

STIC-Biotech/ChemLib

86645

From: McElwain, Elizabeth  
Sent: Wednesday, February 12, 2003 4:28 PM  
To: STIC-Biotech/ChemLib  
Subject: sequence search

Please search for prior art and for interference  
09/856,018 - SEQ ID NO: 15  
and for DNA encoding SEQ ID NO: 16

Thank you,  
Beth

Elizabeth F. McElwain, Ph.D.  
U.S. Patent and Trademark Office  
Tech Center 1600, Art Unit 1638  
room CM1-9E11  
mailbox CM1-9E12  
703-308-1794  
elizabeth.mcelwain@uspto.gov

CRFE

POINT OF CONTACT:  
PAUL SCHULWITZ  
TECHNICAL INFO. SPECIALIST  
CM1 6B06 TEL. (703) 305-1954

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 2/20  
Date Completed: 2/24  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_



Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1821	67.9	1942	4	US-09-326-203A-1	Sequence 1, Appli
2	802.5	29.9	1650	4	US-09-103-754A-2	Sequence 2, Appli
3	800.5	29.9	1976	3	US-09-165-042-2	Sequence 2, Appli
4	784.5	29.3	1766	4	US-09-326-203A-15	Sequence 15, Appli
5	784.5	29.3	1766	4	US-09-326-203A-16	Sequence 16, Appli
6	739.5	27.6	1895	4	US-09-326-203A-14	Sequence 14, Appli
7	568.5	21.2	629	4	US-09-103-754A-3	Sequence 3, Appli
8	396	14.8	4011	1	US-08-121-057-3	Sequence 3, Appli
9	396	14.8	4011	2	US-08-509-187D-3	Sequence 3, Appli
10	396	14.8	4011	2	US-09-121-396-3	Sequence 3, Appli
11	396	14.8	4011	5	PCT-US93-09704A-3	Sequence 3, Appli
12	390	14.6	2040	3	US-09-165-042-4	Sequence 4, Appli

Db 237 ATGGCGATTGGATTCT--GCTGGCGTTACTACGGTGACGGAGAACGGTGGCGGAGAG 293  
Qy 19 -----LeuArgArgProSer-----AlaThrSerThrAlaGlyLeu 31  
Db 294 TTGTCGATCTTGATAGGCTTCGTCGACGGAAATCGAGATCGGATCTTCTAACGGACTT 353  
Qy 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysSerGly 51  
Db 354 CTCTCTCTGCTCCGATAAATAATCTCTCTCGATGATGTGGAGCTCCCGCGCACGCTT 413  
Qy 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67  
Db 414 AGGATCGGATGATTCGCTGTTAAACGATCAGCTCAGGGAACAGCCCAATTTGGCCGGA 473  
Qy 68 GlnAsnGlu-----LysGln 72  
Db 474 GATAATAACGGTGGTGGCGATAATAACGGTGGTGGAGAGCGCGGAGAGAGGA 533  
Qy 73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92  
Db 534 AACCCGATGCTAGC-----TTTACGATATCGACCGCTGCTCCAGCTCATCGGAGG 584  
Qy 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112  
Db 585 CGGAGAGAGAGCTCAGTACCTCGACCAATCTTCAACAGAGCCATGCCGGAATTATTC 644  
Qy 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMet 132  
Db 645 AACCTCTGTAGTAGTCTTATTGCTGTAACAGATAGACTCATCATCGAAAATCTTATG 704  
Qy 133 LysTyrGlyThrLeuLysSerGlyPheThrPheSerSerLysSerLeuArgAspTrp 152  
Db 705 AAGTATGCTGTGATGATCAGAACCGATTCTGTTAGTTCAAGATCGCTCGGAGATTGG 764  
Qy 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172  
Db 765 CCGCTTTTCATGCTGTGTATATCCCTTCGATCTTCTTGGCTGCCCTTACCGTTGAG 824  
Qy 173 LysLeuAlaGlnArgLysCysIleProGluProValValValValLeuHisIleIleIle 192  
Db 825 AAATTGGTACTTCAGAAATACATATCAGAACCTGTGTGTCATCTTCTTCATATTATTATC 884  
Qy 193 ThrSerThrSerLeuPheTyrProValLeuValIleLeuArgCysAspSerAlaPheVal 212  
Db 885 ACCATGACAGAGGTTTGTATPCCAGTTTACCTCACCTAACGCTGATTCGTCTTTTATA 944  
Qy 213 SerGlyValThrLeuMetLeuPheSerCysValValThrLeuLysLeuValSerTyrAla 232  
Db 945 TCAGGTGTCACTTTGATGCTCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 1004  
Qy 233 HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252  
Db 1005 CATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1052  
Qy 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272  
Db 1053 -----AATCCTGAAGTCTCCTACGCTTACGCTTGAAGAGCTTGGCATATTTCATG 1103  
Qy 273 ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGly 292  
Db 1104 GTCCGCTCCACATTTGTGTATCAGCAAGTTATCCAGCTTCTGCTGATATACGGAAGGT 1163  
Qy 293 TrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIle 312  
Db 1164 TGGTGGCTCGTCAATTTGCAAAACTGGTATATATACCGGATTCGGGATTATAATA 1223  
Qy 313 AspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332  
Db 1224 GAACAAATATAAATCTTATGTCAGGAACCTCAAGATCCTTTGAAAGCGGATCTCTTA 1283  
Qy 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352  
Db 1284 TATCCTATTGAACAGCTTGCAAGCTTTCAGTTCCAAATTTATATATGTCGCTCTGCATG 1343

Qy 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372  
Db 1344 TTCTACTCTCTTCCACCTTTGGTTAAACATATTGGCAGAGCTTCTCTGCTTCGGGAT 1403  
Qy 373 ArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392  
Db 1404 CGTGAATCTTCAAAAGATTGGTGGAAATGCAAAAGTGTGGAGATTACTGGAGAAATGTGG 1463  
Qy 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412  
Db 1464 AATATGCTGTTCATAAATGGATGTTCCGACATATATATCTCCCTGCTTCCGACAGCAAG 1523  
Qy 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432  
Db 1524 ATACAAAGACACATCGCCATTATCATCTTCTTAGTCTCTGCAGTCTTTCATGAGCTA 1583  
Qy 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452  
Db 1584 TGCATCGCAGTCTCTTGTCTCTCAAGCTATGGGCTTTTCTTGGGATTTATGTTTCAG 1643  
Qy 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472  
Db 1644 GTGCTTTGGTCTTCATCACAACATCTCTCAGGAAGGTTT---GGCTCAACGGTGGGG 1700  
Qy 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492  
Db 1701 AACATGATCTTCTGTTTCATCTTCTGCATTTTCGACAAACCGATGTTGCTCTCTTTAT 1760  
Qy 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503  
Db 1761 TACCACGACCTGATGAAACCGAAAGGATCGATG 1793

## RESULT 2

US-09-103-754A-2  
; Sequence 2, Application US/09103754A  
; Patent No. 6344548  
; GENERAL INFORMATION:  
; APPLICANT: Farese, Robert  
; APPLICANT: Cases, Sylvaine  
; APPLICANT: Smith, Steven  
; APPLICANT: Erickson, Sandra  
; TITLE OF INVENTION: Diacylglycerol O-acyltran  
; TITLE OF INVENTION: sferase  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bozicevic & Reed  
; STREET: 285 Hamilton Avenue, Suite 200  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/103,754A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Field, Bret E  
; REGISTRATION NUMBER: 37,620  
; REFERENCE/DOCKET NUMBER: 6510-105p  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650 327 3400  
; TELEFAX: 650 327 3231  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:



SEQUENCE CHARACTERISTICS:  
 LENGTH: 1650 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-09-103-754A-2

Alignment Scores:  
 Pred. No.: 1.64e-82  
 Score: 802.50  
 Percent Similarity: 55.16%  
 Best Local Similarity: 37.50%  
 Query Match: 29.94%  
 DB: 4

US-09-856-018B-16 (1-504) x US-09-103-754A-2 (1-1650)

Qy	17	SerSerLeuArgArgArgPro---SerAlaThrSerThrAlaGlyLeuPheAsnSerPro	35
Db	120	AGCTCTCGGCGTCGGAGACGGCTCGCGGTTTCCGTCACGGGT-----	164
Qy	36	GlutThrThrAspSerSerGlyAspLeuAlaValAspSerGlySerAspAspSer	55
Db	165	-----GGTAGTGGCCCCAAGGTACAGAGGACGAGGTCGCGAGACGGGCT	209
Qy	56	IleAsnSerAsp-----AspAlaAlaValAsnSerGlnGlnAsnGlu	70
Db	210	GTGAGCCCGACTGGCGCGGGGTGACGCGCGGCTCCGGCTCCGCGTCCAGCCAT	269
Qy	71	LysGlnAspThrAppPheSerValLeuLysPheAlaIlyArgProSerVal-----	87
Db	270	ACCCGGGACAAAGAC-----GGCGGGACCGAGCGGTGGGGACGGC	308
Qy	88	-----ProAlaHisArgLysValLysGluSerProLeuSerSerAspThrIle	103
Db	309	TACTGGGATCTGAGGTGCCATCGT---CTGCAAGATCTTTGTTCAGCTCAGACAGTGGT	365
Qy	104	PheArgGlnSerHisAlaGlyLeuPheAsnLeuCysIleValValLeuValAlaValAsn	123
Db	366	TTC---AGCAATATCTGGGTATCTGAATTGGTGTGGTGATCTGATCTGAGTAAT	422
Qy	124	SerArgLeuIleLeuGluAsnLeuMetLysTyrGlyTrpLeuIleLysSerGlyPheTrp	143
Db	423	GCAAGGTATTTTAGAAGACCTATCAAGATATGGCATCTCGTGGT---GATCCTATCCAG	479
Qy	144	PheSerLysSerLeuArgAsp-----TrpProLeuPheMetCysCysLeuSer	160
Db	480	GTGGTGTCTCTGTTTTGAAGGACCCCTACAGCTGGCTGCCCATGGCTGATATTGCA	539
Qy	161	LeuValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIle	180
Db	540	TCCAATATTTTGTGTGGCTGCATCTCAGATTGAGAACGCCCTGGCAGTGGGTGCCCTG	599
Qy	181	ProGluProValValValLeuHisIleIleIleThrSerThrSerLeuPheTrpPro	200
Db	600	ACAGACAGATGGGGCTGCTCATCTACATGTGGTTAACCTGGCCCAATCATTTGCTCCCA	659
Qy	201	ValLeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPhe	220
Db	560	GCACCTGTGGCTTACTGGTTGATGCTATCTACTCCAGTGGGTTCGGTGTTCGCTCGCA	719
Qy	221	Ser---CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyr-----	236
Db	720	TCATACCTCATGTTCTTCAAGCTTATTCCTACCGGGATGTCACACCTGGTGGCGC	779
Qy	237	AspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeu---	255
Db	780	CAGCGAAGGTCGAGGCAAGGCTGCTCTACAGGGAAGAAGGTCAAGTGGGCTGCTGCC	839
Qy	256	-----AsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuVal	273
Db	840	CAGCAAGCTGTAGCTATCAGACAACTGACCTACCTACCGAGATCTCTATTCATCTTT	899





Db 225 AAAGACCCGACAGCAGCGCTGGCGGACGCCACTGG-----GAGCTGAGGTGCCATCGT 278  
Qy 92 LysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeu 111  
Db 279 ---CTGCAAGACTCTTGTTCAGCTCAGCAGACGGGTTTC---AGCAATTACCGTGGTATC 332  
Qy 112 PheAsnLeuCysIleValValValValAlaValAsnSerArgLeuIleGluAsnLeu 131  
Db 333 CTGAATTGGTGGTGGTGTGATGCTGATCCTCGAGTAAAGTCAAGGTATTTTATAGAAATCTT 392  
Qy 132 MetLysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAsp 151  
Db 393 ATCAAGTATGGCATCCTGTGTG---GATCCCATCCAGGTGGTGTCTCTCTTCTGAAGGAC 449  
Qy 152 -----TrpProLeuPheMetCysLysLeuSerLeuValValPheProPheAlaLa 168  
Db 450 CCTACAGCTGGCTGCCCTCCCTCATGATGATGCATCAATATCTTTATTGTGGCTACA 509  
Qy 169 PheIleValGluLysLeuAlaGlnArgLysCysIleProGluProValValValValLeu 188  
Db 510 TTTCAAGATTGAGAAGCGCTGTCACTGGTGGCTGCCTGACAGACAGATGGGCTGCTGCTA 569  
Qy 189 HisIleIleIleThrSerThrSerLeuPheTrpProValLeuValIleLeuArgCysAsp 208  
Db 570 CATGTGGTTAACTGGCCACAATATCTCTCTCCAGCAGCTGTGGCTTACTGTGTGAG 629  
Qy 209 SerAlaPheValSerGlyValThrLeuMetLeuPheSer---CysValValTrpLeuLys 227  
Db 630 TCTATCACCTCCAGTGGTTCCTGTGCTGTGGCATATCTCCATCATCTTCCTCAG 689  
Qy 228 LeuValSerTyrAlaHisThrAsnTyr-----AspMetArgAlaLeuThrLysLeu 244  
Db 690 CTTTCTCTACCGGATGCTCAATCTGTGGTGGCGCAGCAGGAGGTCAAGCCAAAGCT 749  
Qy 245 ValGluLysGlyGluAlaLeuLeuAspThrLeu-----AsnMetAspTyrProTyr 261  
Db 750 GTGTCTGCGAGGAAGGATGAGTGGGCTGCTGCCCAAGCACTGTAAGCTATCCGGAC 809  
Qy 262 AsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGlnPro 281  
Db 810 AACCTGACCTACCGAGATCTCTATTACTTCACTCTTGTCTCTTCTTGTGTTTGAACCTC 869  
Qy 282 SerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeuValLysLeu 301  
Db 870 AACTTTCCTCGATCCCGCAATACAAAGCGCTTCTGCTACGGCGGTTCTTTCAGATG 929  
Qy 302 IleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIleValGln 321  
Db 930 CTCCTTTTCACCCAGCTTCAAGTGGGCTGATCCAGCAGTGGATGCTCTACTATCCAG 989  
Qy 322 AsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAla-----ThrGluArgValLeu 339  
Db 990 AACTCCATGAAGCCCTTCAAG---GACATGGAGCTATTTCAGCAATCTTTCAGCGTCTCTTA 1046  
Qy 340 LysLeuSerValProAsnLeuTyrValTrpLeuLysMetPheTyrCysPhePheHisLeu 359  
Db 1047 AAGCTGGCGGTCCCAACCATCTGATATGGCTCATCTTCTTCTTATGGCTTTTCCACTCA 1106  
Qy 360 TrpLeuAsnIleLeuAlaGluLeuArgPheGlyAspArgGluPheTyrLysAspTrp 379  
Db 1107 TGTCTCAATGCTGGCAGAGCTCTGCAAGTTGGAGACCGCGAGTTCTACAGGAGCTGG 1166  
Qy 380 TrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTrp 399  
Db 1167 TGAATGTGTGACTGTCTACCTTCTTTTGGCAGAACTGGAATATCCCGTGCACAGTGG 1226  
Qy 400 MetIleArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAlaLeu 419  
Db 1227 TGCATCAGACACTTCTACAAGCTTGTCTCAGCTGGCGCAGCAACAATGGATGCCAGG 1286  
Qy 420 LeuIleAlaPheLeuValSerAlaLeuPheHisGluLeuCysIleAlaValProCysHis 439  
Db 1287 ACTGGGCTCTTTTGGCTCAGCCCTTCTTCCATGAGTACCTAGAGCATTCCTCCCTGAGG 1346

Qy 440 IlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThr 459  
Db 1347 ATGTTCCGCTCTGGCATTCACAGCATGATGGCTCAGTCCCATGGCC-----1397  
Qy 460 AsnTyrLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIle 479  
Db 1398 ---TGCATTGTGAACCGCTTCTTCCAAGGAACATATGCAATGCAGCTGTGTGG---GTG 1451  
Qy 480 PheSerIleLeuGlyGlnProMetCysValLeuLeuTyrTyrHisAsp 495  
Db 1452 ACACATCATTTGGGCAACCGGTGGCTGTGCTCATGTATGTCCACGAC 1499  
RESULT 6  
US-09-326-203A-14  
; Sequence 14, Application US/09326203A  
; Patent No. 6444876  
; GENERAL INFORMATION:  
; APPLICANT: Lassner, Mike  
; APPLICANT: Ruzinsky, Diane  
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
; FILE REFERENCE: 17045/00/WO  
; CURRENT APPLICATION NUMBER: US/09/326,203A  
; PRIOR FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 60/088,143  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/108,389  
; PRIOR FILING DATE: 1998-11-12  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 1895  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (209)  
; OTHER INFORMATION: n at position 209 is unknown  
US-09-326-203A-14

Alignment Scores:  
Pred. No.: 3,78e-75 Length: 1895  
Score: 739.50 Matches: 158  
Percent Similarity: 59.53% Conservative: 70  
Best Local Similarity: 41.25% Mismatches: 138  
Query Match: 27.59% Indels: 17  
DB: 4 Gaps: 9

US-09-856-018B-16 (1-504) x US-09-326-203A-14 (1-1895)

Qy 123 AsnSerArgLeuIleIleGluAsnLeuMetLysTyrGlyTrpLeuIleLysSerGlyPhe 142  
Db 243 AATGCCGGTATTCTTCTGGAAGACCTCATCAAGTATGGCATCTCTGGTG---GACCCCATC 299  
Qy 143 TrpPheSerSerLysSerLeuArgAsp-----TrpProLeuPheMetCysLysLeu 159  
Db 300 CAGGTGGTTTCTCTGTCTCTGAAGGATCCCTATAGTGGCCGCCCATGCTGGTTATT 359  
Qy 160 SerLeuValValPheProPheAlaPheIleValGluLysLeuAlaGlnArgLysCys 179  
Db 360 GCGGCAATGTCTTCTGCTGTGCTGCATTCCAGTTTGAAGCGCTCGGCTGGTGCC 419  
Qy 180 IleProGluProValValValValLeuHisIleIleIleThrSerThrSerLeuPheTyr 199  
Db 420 CTGACGGACGGGAGCTGCTGTGACGTGGCCCACTGGCCCACTATCTGTGTTTC 479  
Qy 200 ProValLeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeu 219  
Db 480 CCAGCGGCTGTGTTCTTACTGTTGAGTCTATCACTCCAGTGGGCTCCCTGCTGGCGCTG 539  
Qy 220 PheSer---CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsn---TyrAsp 237  
Db 220 PheSer---CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsn---TyrAsp 237

Db 540 ATGGCGCACACCATCTCTCTCAAGCTCTTCTCTACCCGAGCTCAACTCATGGTGC 599  
Qy 238 MetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeu----- 255  
Db 600 CGCAGGCGCAGGCGCAGGCTCTCTCGACGGAAGAGCGCAGCTGCTGCTGCCCG 659  
Qy 256 ---AsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAla 274  
Db 660 CACACCGTGAGTACCCGGAATCTGACCTACCGGATCTCTACTCTCTCTCTGCCC 719  
Qy 275 ProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeu 294  
Db 720 CCCACCTTGCTACGAGCTCAACTTCCCGCTCTCCCGCATCCGGAACGCTTCTG 779  
Qy 295 PheArgGlnLeuValLysLeuIlePhePheThrGlyValMetGlyPheIleIleAspGln 314  
Db 780 CTGGAGGGATCTTGAGTCTGCTTCTACCCAGCTCCAGGTGGGGCTGATCCAGCAG 839  
Qy 315 TyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAla 334  
Db 840 TGGATGTGCCACACCACTCCAGACTCCATGAGCCCTTCAAG---GACATGGACTACTCA 896  
Qy 335 -----ThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352  
Db 897 CGCATCATCGAGCGCTCTGAACTGGCGTCCCAATCACTCATCTGCTCATCTTC 956  
Qy 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372  
Db 957 TTCCTACTGGCTCTCCACTCTGCTGAATGCCGTGAGCTCATGCAGCTTTGGAGAC 1016  
Qy 373 ArgGluPheTyrLysAspTyrTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392  
Db 1017 CGGGAGTCTACCGGGACTGTGTGAATCCGAGTCTGTCACTACTCTTGTGCAAACTGG 1076  
Qy 393 AsnMetProValHisLysTyrMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412  
Db 1077 AACATCCTGTGCACAGTGTGTGATCAGACATCTTACAGCCCATCTTCGAGGGGC 1136  
Qy 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432  
Db 1137 AGCAGCAAGTGGATGGCGCAGGACGAGGGGTCTCTGGCGCTTCTTCCACGAGTAC 1196  
Qy 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452  
Db 1197 CTGTGAGCGTCCCTTCGGAATCTTCGCGCTCTGGCGTTCAGGGCATGATGGCTCAG 1256  
Qy 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472  
Db 1257 ATCCCACTGGCC-----TGGTCTGGGCGCGCTTTTCCAGGGCAACTATGGC 1304  
Qy 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492  
Db 1305 AACGAGCTGTGTGG---CTGCTGCTCATCATCGGACGACCAATAGCGCTCTCATGTAC 1361  
Qy 493 TyrHisasp 495  
Db 1362 GTCCACGAC 1370

## RESULT 7

US-09-103-754A-3

Sequence 3, Application US/09103754A

Patent No. 634548

GENERAL INFORMATION:

APPLICANT: Farese, Robert

APPLICANT: Cases, Sylvaine

APPLICANT: Smith, Steven

APPLICANT: Erickson, Sandra

TITLE OF INVENTION: Diacylglycerol O-acyltran

TITLE OF INVENTION: sferase

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bozicevic &amp; Reed

STREET: 285 Hamilton Avenue, Suite 200

RESULT 8

US-08-121-057-3

CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/103.754A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Field, Bret E  
REGISTRATION NUMBER: 37,620  
REFERENCE/DOCKET NUMBER: 6510-105p  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650 327 3400  
TELEFAX: 650 327 3231  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 629 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-103-754A-3

Alignment Scores:

Pred. No.: 3,2e-56 Length: 629

Score: 568.50 Matches: 105

Percent Similarity: 87.02% Conservative: 9

Best Local Similarity: 80.15% Mismatches: 13

Query Match: 21.21% Indels: 4

DB: 4 Gaps: 1

US-09-856-018B-16 (1-504) x US-09-103-754A-3 (1-629)

Qy 289 IleArgLysGlyTrpLeuPheArgGlnLeuValLysLeuIleLeuPheThrGlyValMet 308  
Db 8 ATACGGAAGGTTGGTGGCTCGTCAATTGCAAACTGGTCAATATTCACCGGATTCATG 67  
Qy 309 GlyPheIleIleAspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLys 328  
Db 68 GGATTTATATAGAACAAATATATAATCTATTGTGAGAACTCAAGCATCTCTTTGAAA 127  
Qy 329 GlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrVal 348  
Db 128 GGCGATCTCTATATGCTATTGAAAGAGTGTGAAGCTTTTCAGTTCCAAATTTATATGTG 187  
Qy 349 TrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeu 368  
Db 188 TGGCTCTGCTACTTCTACTGCTTCTTCCACCTTTGGTTAAACATATTCGACAGGCTTCTC 247  
Qy 369 ArgPheGlyAspArgGluPheTyrLysAspTrpAsnAlaLysThrValGluAspTyr 388  
Db 248 TGCTTCGGGGATCGTGAATCTACAAAGATTGGTGAATGCAAAAGTGTGGGAGATTAC 307  
Qy 389 Trp-ArgMetTrpAsnMetProValHisLysTrp-MetIleArgHisLeuTyrPhePro 407  
Db 308 TGGGAGAATGTGGAATATGCTGTCCATAATGGGATGGGTCGCGACATATATACCTTCCC 367  
Qy 408 ---CysLeuArgHisGlyLeuProLys 415  
Db 368 CGTGTTCGCGCACAGGATTACCCAAA 394

Sequence 3, Application US/08121057  
 Patent No. 5484727  
 GENERAL INFORMATION:  
 APPLICANT: CHANG, TA-YUAN  
 APPLICANT: CHANG, CATHERINE C. Y.  
 TITLE OF INVENTION: ACYL COENZYME A: CHOLESTEROL  
 TITLE OF INVENTION: ACYLTRANSFERASE (ACAT)  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 60 STATE STREET, SUITE 510  
 CITY: BOSTON  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/121,057  
 FILING DATE:  
 CLASSIFICATION: 800  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LAMPORT HAMMITTE, ANN.  
 REGISTRATION NUMBER: 34,858  
 REFERENCE/DOCKET NUMBER: DCI-033CP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-2700  
 TELEFAX: (617) 227-5941  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4011 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1397..3046  
 US-08-121-057-3

Alignment Scores:  
 Pred. No.: 5,21e-35 Length: 4011  
 Score: 396.00 Matches: 150  
 Percent Similarity: 44.66% Conservative: 80  
 Best Local Similarity: 29.13% Mismatches: 191  
 Query Match: 14.78% Indels: 95  
 DB: 1 Gaps: 24

US-09-856-018b-16 (1-504) x US-08-121-057-3 (1-4011)

QY 33 AnSerProGluThrThrThrAspSerSerGly-AspAspLeu-----AlaLy 48  
 Db 1537 GATGCAAGAGATAAAGATTGACGAGCAGAGAGGAAATTTATGAA 1596  
 QY 48 sAspSerGlySerAspSerIleAsnSerAspAlaAlaValAsnSerGlnGlnG 68  
 Db 1597 GGAAGTTGGCAGT-----CATTGTGATTTTGTGACCAATCTCATTTGAAA 1644  
 QY 68 nAsnGluLysGlnAsp-----ThrAspPheSerValLeu----- 79  
 Db 1645 GTCAGCATCATTAGATAAATGGTGGGCGCTCTCACAACTTTCTGTCTTGAAGAGA 1704  
 QY 80 -LysPheAlaTyrArgProSer-----ValPro----- 88  
 Db 1705 GAAACACACCACTGAGCGGAGGATTGAGAGCACCCTCCAGAACAGGAAAGATTTTAT 1764  
 QY 89 -AlaHisArgLysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerH 108  
 Db 1765 TGCAGGCCCTCTCTTATAGTAGACTGCTTGAAGTGGACCATC----- 1810

QY 108 sAlaGlyLeuPheAsnLeuCysIleValValLeuValAlaA---ValAsnSerArgLeu 127  
 Db 1811 -AGACAATATATACATGTTTATGCTCCCTCCTCATCTCTTATCCTCAGCACACATGT 1869  
 QY 127 eileGlu-AsnLeuMetLysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerL 147  
 Db 1870 AGTAGATTACATTGATGAAGAGGCTGGTGTGAGT-----TCAGCCT 1914  
 QY 147 ysSerLeuArg-AspTrpProLeuPheMetCysCysLeuSerLeuVal----- 162  
 Db 1915 CTGCTCTATGCTTTTGGCAAAATTTCTACCGTTGTTTGGACCTGGATCATGTTTCC 1974  
 QY 163 ValPheProPheAlaAla---PheIleValGluLysLeuAla-----GlnArg 177  
 Db 1975 GTCTACATTTTTCAGTTCCCTATTTTCTGTTTCAACATTTGGCGCACCTGGCTATAGCAAG 2034  
 QY 178 LysCysIleProGluProValVal-----ValValLeuHisIleIleIleThrSer 194  
 Db 2035 TTCTCATCCGCTGATCCGTTCTCTCTTCCATGGCTTTCTTTTCATGATCTTCCAGATTG 2094  
 QY 195 ThrSerLeuPheTyr-ProValLeuValIleLeuArgCysAspSerAlaPheValSerG 214  
 Db 2095 AGTTCTAGCTTTTGGACCAACATATGTTGTGTTAGCATATACACTCCACACCTTCCCG 2154  
 QY 214 yValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAlaHis 234  
 Db 2155 GTTCATCATATATATTCGACAGATTCGTTTGTGTAATGAAGGCCCACTCATTTGTCAGAGA 2214  
 QY 234 rAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeu---LeuAs 253  
 Db 2215 GAACGTGCCT---CGGGTACTAAATTCAGCTAAGAGAGAAATCAAGCACTGTTCCAAATAC 2271  
 QY 253 pThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuVa 273  
 Db 2272 TACAGTCAAC-----CAGTATTGTACTTCTTATT 2301  
 QY 273 lAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTr 293  
 Db 2302 TGCTCTCATCCCTTATCTACCGTGACAGCTATCCAGGAATCCACACTGAAGATGGGTTA 2361  
 QY 293 pleuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleLeas 313  
 Db 2362 TGTGCTATGAAGTTTGCACAGCTCTTGTGTTGCTTTTCTATGTGTACTATCATCTTTGA 2421  
 QY 313 pGlnTyrIleAsnProIleValGlnAsn---SerGlnHisProLeuLysGlyAsnLeuLe 332  
 Db 2422 AAGGCTTTGTGCCCTTGTTCGGAAATATCAACAGAGAGCCCTTCAGCGCTCGTTCT 2481  
 QY 332 uTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMe 352  
 Db 2482 GGTCTCTATCT-----GTATTTAACTCCATCTTGCCAGGTGCTCATCTCTCTCTTAC 2535  
 QY 352 tPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuArgPheGlyAs 372  
 Db 2536 TTTTGTGCTTTTGTGCTGCTGCTCAATGCTTTTGTGAGATGTTACGCTTTGGTGA 2595  
 QY 372 pArgGluPheLysAspTrpAsnAlaLysThrValGluAspTyrTrpArgMetTr 392  
 Db 2596 CAGGATGTTCTATAGGATTGGTGGAACTCCAGCTCATCTACTTCAACTATTATAGAACCTG 2655  
 QY 392 pAsnMetProValHisLysTrpMetIleArgHis-----LeuTyrPhePr 407  
 Db 2656 GAATGGTGGTCCATGACTGGCTATATATGCTTACAAGGACTTCTCTGGTTTTT 2715  
 QY 407 cCysLeuArgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAl 427  
 Db 2716 CTCCAAGAGATTC-----AAATCTGCTGCCATGTTAGTGTCTTCTGCTGTATCTGC 2766  
 QY 427 aleuPheHisGluLeuCysIleAlaValProCys-----HisIlePh 441  
 Db 2767 TGTAGTACAGGAATATGCTTGTGCTGTT---TCTTACAGCTTTTCTATCCCGTCTGTT 2823  
 QY 441 eLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThrAsnTy 461

Db 2824 CAGTCTCTTCATGTTCTTGGAAATGGCTTC-----ACCTT 2859  
Qy 461 rLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSe 481  
Db 2860 CATTCGAATGATAGTCGGAAGAAAGCCGATTTGGAATGTTCTGATGTGGAC---TTCTC 2915  
Qy 481 rIleLeuGlyGlnPrometCysValLeuLeuTyrTyr 493  
Db 2916 TTTTCTGGCAATGGAGTCTTACTCTGCTTTTATTC 2952

RESULT 9  
US-08-509-187D-3  
; Sequence 3, Application US/08509187D  
; Patent No. 5834283  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Ta-Yuan and Chang, Catherine C.Y.  
; TITLE OF INVENTION: ACYL Coenzyme A: Cholesterol Acyltransferase  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/509,187D  
; FILING DATE: 31-JUL-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lampert Hammitte, Ann  
; REGISTRATION NUMBER: 34,858  
; REFERENCE/DOCKET NUMBER: DCI-0333cpdv  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4011 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1397..3046  
US-08-509-187D-3

Alignment Scores:  
Pred. No.: 5,21e-35 Length: 4011  
Score: 396.00 Matches: 150  
Percent Similarity: 44.66% Conservative: 80  
Best Local Similarity: 29.13% Mismatches: 191  
Query Match: 14.78% Indels: 95  
DB: 2 Gaps: 24

US-09-856-018B-16 (1-504) x US-08-509-187D-3 (1-4011)  
Qy 33 AsnSerProGluThrThrThrAspSerSerGly-AspAspLeu-----AlaIy 48  
Db 1537 GATGCAAGAAAGATAAAGTTGACAGCAGAGCAGAGGAATTCAGACCCATTTTATGAA 1596  
Qy 48 sAspSerGlySerAspSerIleAsnSerAspAspAlaAlaValAsnSerGlnGlnGln 68

Db 1597 GGAAGTTGGCAGT-----CACTTTGATGATTTTGTGACCAATCTCATTTGAAA 1644  
Qy 68 nAsnGluLysGlnAsp-----ThrAspPheSerValLeu----- 79  
Db 1645 GTCAGCATCATATAGATAATGGTGGTGGCTCTCAACACCTTTTCTGTCTTGAAGAGA 1704  
Qy 80 -LysPheAlaTyrArgProSer-----ValPro----- 88  
Db 1705 GAAAAACAACCATGACGGAAGGATTTGAGACACCTCCAGAACAGGAAGATTTTAT 1764  
Qy 89 -AlaHisArgLysValLysGluSerProLeuSerSerSerThrIlePheArgGlnSerHi 108  
Db 1765 TGCAGGGCGCTCTCTTAGATGAACCTGCTGAAGTGGACACATC----- 1810  
Qy 108 sAlaGlyLeuPheAsnLeuCysIleValValLeuValAla-----ValAsnSerArgLeuI 127  
Db 1811 -AGAACAATATATACATGTTTATGCGCTCTCATCTCTTTATCTCCTCAGCACACTTGT 1869  
Qy 127 eIleGlu-AsnLeuMetLysTyrGlyTyrLeuIleLysSerGlyPheTrpPheSerSerL 147  
Db 1870 AGTAGATTACATTGATGAAGGAGGCTGGTCTTGAGT-----TCAGCCT 1914  
Qy 147 ysSerLeuArg-AspTrpProLeuPheMetCysCysLeuSerLeuVal----- 162  
Db 1915 CCTGCTTATGCTTTGGCAATTTCTTACCGCTTGTGGACCTGGTGGATCATGTTCT 1974  
Qy 163 ValPheProPheAlaAla---PheIleValGluLysLeuAla-----GlnArg 177  
Db 1975 GTCTACATTTTTCAGTCTCCCTATTTCTGTTTCAACATTTGGCGCAGCTGGTATAGCAAG 2034  
Qy 178 LysCysIleProGluProValVal-----ValValLeuHisIleIleIleIlehrSer 194  
Db 2035 TTCTCATCGCTGATCGGTTCTCTTCCATGGCTTTCTTTTCATGATCTCTCCAGATTGG 2094  
Qy 195 ThrSerLeuPheTyr-ProValLeuValIleLeuArgCysAspSerAlaPheValSerG 214  
Db 2095 AGTTCTAGGTTTGGACCAACATATGTTGTGTAGCATATACATGCCACCACTTCCG 2154  
Qy 214 yValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAlaHisTh 234  
Db 2155 GTTCATCATATATTCGAGCAGATCGTTTGTGTAATGAAGGCCACTCATTTGTGAGAGA 2214  
Qy 234 rAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeu---LeuAs 253  
Db 2215 GAACGTGCTGCT---CGGGTACTAAATTCAGCTAAGGAGAAATCAAGCAGCTGTTCCAATACC 2271  
Qy 253 pThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuVa 273  
Db 2272 TACAGTCAAC-----CAGTATTCTTCTTCTTAT 2301  
Qy 273 lAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTr 293  
Db 2302 TGCTCCTACCTTATCTACCGTGACAGCTATCCAGGAATCCCACTGTANGATGGGTTA 2361  
Qy 293 pLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIleAs 313  
Db 2362 TCTCGCTATGAAGTTGACAGGCTCTTTGGTGTCTTTTCTATGTGTACTACTCTTTGA 2421  
Qy 313 pGlnTyrIleAsnProIleValGlnAsn---SerGlnHisProLeuLysGlyAsnLeuLe 332  
Db 2422 AAGGCTTTTGCGCCCTTGTTCGGAATATCAACAGGAGCCCTTCAGCGCTCGTGTCT 2481  
Qy 332 uTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMe 352  
Db 2482 GGTCTCTATGT-----GTATTTAACTTCCACTCTTGCAGGTGTGCTGATTCTCTTCTAC 2535  
Qy 352 tPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAs 372  
Db 2536 TTTTGTGCTTTTTCGACTGCTGCTCAATGCCCTTGTGAGATGTTTACGCTTTGGTGA 2595  
Qy 372 pArgGluPheTyrLysAspTyrTrpAsnAlaLysThrValGluAspTyrTrpArgMetTr 392  
Db 2596 CAGGATGTTCTATAGGATTTGGTGAACCTCCAGCTACTCCAACTATATATAGAACCTG 2655







```
Db 2422 AAGGCTTTGTGCCCTTGTTCGGAATATCAACAGAGAGCCCTTCAGCGCTCGTTCT 2481
Qy 332 uTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMe 352
Db 2482 GGTCTCTATGT-----GTATTAACTCCATCTTGCAGGTGTGCTGATTCCTCTTAC 2535
Qy 352 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAs 372
Db 2536 TTTTTCCTTTCCTTTCGCTGCTGCTCAATGCCCTTGTCTGAGATGTTACGCTTTGGTGA 2595
Qy 372 PArgGluPheTyrLysAspTrpPheAsnAlaLysThrValGluAspTrpTyrPheMetTr 392
Db 2596 CAGGATGTTCTATAGAGATTGGTGAACCTCCAGCTCATCTACTCAACTATTATAGAACCTG 2655
Qy 392 PAsnMetProValHisLysTrpMetIleArgHis-----LeuTyrPhePr 407
Db 2656 GAATGTGCTGCCATGCTGCTATATCTACTATCTTACAAAGGACTTTCCTGCTGTTT 2715
Qy 407 cCysLeuArgHisGlyLeuProLysAlaAlaAlaLeuIleAlaPheLeuValSerAl 427
Db 2716 CTCCAAGAGATTC-----AAATCTGCTGCCATGTAGTCTTTGCTGTATCTGC 2766
Qy 427 AleuPheHisGluLeuCysIleAlaValProCys-----HisIlePh 441
Db 2767 TGTACTACAGATATGCTTGGCTGTT--TGCTTGAGCTTTTCTATCCGCTGCTGT 2823
Qy 441 eLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThrAsnTy 461
Db 2824 CTGCTCTTCATGTTCTTGGAAATGGCTTTC-----AACTT 2859
Qy 461 rLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSe 481
Db 2860 CATTCTCAATGATAGTCGGAAGAAAGCCGATTGGAAATGTTCTGATGGAC-----TTCTC 2915
Qy 481 rIleLeuGlyGlnProMetCysValLeuLeuTyrTrp 493
Db 2916 TTTTCTTGGCAATGAGCTTACTCTGCTTTATTC 2952

RESULT 11
PCT-US93-09704A-3
; Sequence 3, Application PC/TUS9309704A
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: ACYL COENZYME A: CHOLESTEROL
; TITLE OF INVENTION: ACYLTRANSFERASE (ACAT)
; NUMBER OF SEQUENCES: 9
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/09704A
; FILING DATE: October 12, 1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. SER. NO. 959,950
; FILING DATE: October 14, 1992
; APPLICATION NUMBER: U.S. SER. NO. 121,057
; FILING DATE: September 10, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LAMPORT HAMMITTE, ANN.
; REGISTRATION NUMBER: 34,858
; REFERENCE/DOCKET NUMBER: DCI-033CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-2700
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: CDNA
PCT-US93-09704A-3
Alignment Scores:
Pred. No.: 5,21e-35 Length: 4011
Score: 396.00 Matches: 150
Percent Similarity: 44.66% Conservative: 80
Best Local Similarity: 29.13% Mismatches: 191
Query Match: 14.78% Indels: 95
DB: 5 Gaps: 24

US-09-856-018B-16 (1-504) x PCT-US93-09704A-3 (1-4011)
Qy 33 AsnSerProGluThrThrThrAspSerSerGly-AspAspLeu-----AlaLys 48
Db 1537 GATACAAAGAAGATAAAAGTTTCACAGCAGAGAGCAGAGAAATGAAGCCATTTTATGAA 1596
Qy 48 sAspSerGlySerAspAspSerIleAsnSerAspAspAlaAlaValAsnSerGlnGlnG 68
Db 1597 GGAAGTTGGCAGT-----CACTTTGATGATTTTGTGACCAATCTCATTTGAANA 1644
Qy 68 nAsnGluLysGlnAsp-----ThrAspPheSerValLeu----- 79
Db 1645 GTCAGCATCATTAGATAGTAATGGTGGCGCTCTCACACCTTTTCTGTCTTGAGGAGA 1704
Qy 80 -LysPheAlaTyrArgProSer-----ValPro----- 88
Db 1705 GAAACAAACCATAGACGCAAGGATTTGAGAGCACCTCCAGAACAGAAAGATTTTAT 1764
Qy 89 -AlaHisArgLysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHi 108
Db 1765 TGCAAGGCGCTCTCTCTAGATGAACCTGCTTGAAGTGAGCACATC----- 1810
Qy 108 sAlaGlyLeuPheAsnLeuCysIleValValLeuValAla---ValAsnSerArgLeuI 127
Db 1811 -AGAACATATATACATGTTTATTTGCCCTCTCATTTCTTTATCTCAGCACACTGT 1869
Qy 127 eIleGlu-AsnLeuMetLysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerL 147
Db 1870 AGTAGATTACATTGATGAAGGAAGCTGCTGCTGAGT-----TCAGCCT 1914
Qy 147 ySerLeuArg-AspTrpProLeuPheMetCysLysLeuSerLeuVal----- 162
Db 1915 CTGTCTTATGCTTTGGCAAAATTTCCCTACCGTGTGGACCTGGTGATCATGTTCTCT 1974
Qy 163 ValPheProPheAlaAla---PheIleValGluLysLeuAla-----GlnArg 177
Db 1975 GTCTACATTTTCAGTTCCTATTTTCTGTTCACATTTGGCGCACTGGCTATAGCAAGAG 2034
Qy 178 LysCysIleProGluProValVal-----ValValLeuHisIleIleIleThrSer 194
Db 2035 TTCTCATCCGCTGATCCGTTCTCTCTCCATGGCTTTCTTTTCATGATGATCTTCCAGATTGG 2094
Qy 195 ThrSerLeuPheTyr-ProValLeuValIleLeuArgCysAspSerAlaPheValSerGl 214
Db 2095 AGTCTAGTGTGGACCAACATATGTTGTGTAGCATATACATGCCACCGCTTCCG 2154
Qy 214 yValThrLeuMetLeuPheSerCysValTrpLeuLysLeuValSerTyrAlaHisTh 234
Db 2155 GTTCATCATTTATATTCGAGCAGATTCGTTTGTGTAATGAAGGCCCACTCATTTGTCCAGAGA 2214
Qy 234 rAsnTyrAspMetArgAlaLeuThrLysLeuValGlyGlyGluAlaLeu---LeuAs 253
Db 2215 GAACGTGCTCT---CGGGTACTAAATTCAGCTAAAGGAGAAATCAAGCACTGTTTCCAATACC 2271
Qy 253 pThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuVa 273
Db 2272 TACAGTCAAC-----CAGTATTGCTACTTCTTAT 2301
Qy 273 lAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTr 293
Db 2302 TGCTCTACCTTATCTACCGTGACAGCTATCCAGGAATCCACCTGTAAAGATGGGGTTA 2361
```

QY 293 pLeuPheArgGlnLeuValLysLeuLeuLeuPheThrGlyValMetGlyPheLeuLeuLeu 313  
Db 2362 TGTGCGTATGAAGTTTGCACAGCTTTTGGTGTCTTTTCTATGTACTATCATCTTTGA 2421  
QY 313 pGlnTyIleAsnProIleValGlnAsn---SerGlnHisProLeuLysGlyAsnLeuLe 332  
Db 2422 AAGCGTTTGGCCCTTTTGGTGTCTTTCGGAATATCAAAACAGGAGCCCTTCAGCGCTCGTGTCT 2481  
QY 332 uTyraIleThrGluArgValLeuLysLeuSerValProAsnLeuTyIleValTrpLeuCysMe 352  
Db 2482 GGTCTCTATGT-----GTATTTAACTCCATCTTGGCAGGTGTCTGTCTCTCTTAC 2535  
QY 352 tPheTyIleCysPhePheHisLeuTrpLeuAsnLeuLeuAlaGluLeuLeuArgPheGlyAs 372  
Db 2536 TTTTGTGCTTTTGCCTTTTGCACCTGCTGCTGCTCAATGCTTTTGTCTGAGATGTTACGCTTTGGTGA 2595  
QY 372 pArgGluPheTyIleAsnTrpIleAsnAlaLysThrValGluAspTyIleTrpArgMetTr 392  
Db 2596 CAGGATGTTCTATAGGATGTTGGAACTCCACGTATATCACTATATAGAACCTG 2655  
QY 392 pAsnMetProValHisLysTrpMetIleArgHis-----LeuTyIlePhePr 407  
Db 2656 GAATGCTGCTCCACTGCTGCTATATCTATCTTACAGGACTTTCTCTGCTTTT 2715  
QY 407 oCysLeuArgHisGlyLeuProLysAlaAlaLeuLeuLeuAlaPheLeuValSerAl 427  
Db 2716 CTCAAAGAGATTC-----AAATCTGCTGCATGTTAGCTGTCTTGTCTATCTGC 2766  
QY 427 aLeuPheHisGluLeuCysIleAlaValProCys-----HisIlePh 441  
Db 2767 TGAGTACAGATATGCTGCTGCTGTT---TGCTGAGCTTTTCTATCCCGTGTCTGT 2823  
QY 441 eLysLeuTrpAlaPheGlyIleMetPheGlnValProLeuValLeuIleThrAsnTy 461  
Db 2824 CGTCTCTTCTATCTTCTTGGATGCTTTC-----AACIT 2859  
QY 461 rLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSe 481  
Db 2860 CATTTGCAATGATAGTACGGAAGGCGGATTTGGAATGTTGATGTGGAC-----TTCTC 2915  
QY 481 rLeuGlyGlnProMetCysValLeuLeuTyIle 493  
Db 2916 TTTTCTGGCAATGAGTCTTACTCTGCTTTTATTC 2952

## RESULT 12

US-09-165-042-4  
; Sequence 4, Application US/09165042  
; Patent No. 6100077  
; GENERAL INFORMATION:  
; APPLICANT: Sturley, Stephen L.  
; APPLICANT: Oskers, Peter  
; TITLE OF INVENTION: ISOLATION OF A GENE ENCODING DIACYLGLYCEROL  
; TITLE OF INVENTION: ACYLTRANSFERASE  
; FILE REFERENCE: 0575/56331  
; CURRENT APPLICATION NUMBER: US/09/165,042  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 2040  
; TYPE: DNA  
; ORGANISM: Yeast  
US-09-165-042-4

## Alignment Scores:

Pred. No.:	8,44e-35	Length:	2040
Score:	350.00	Matches:	143
Percent Similarity:	42.73%	Conservative:	92
Best Local Similarity:	26.00%	Mismatches:	202
Query Match:	14.55%	Indels:	113
DB:	3	Gaps:	26

US-09-856-018b-16 (1-504) x US-09-165-042-4 (1-2040)

QY 16 HisSerSerLeuArgArgProSerAlaThrSer-----ThrAlaGlyLeuPheAsnSe 34  
Db 48 CACCATGGAGCCAGCGCGGCGCTCTGCTGCTGCAGAGACAGAGGCTG-----GG 101  
QY 34 rProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGlySerAspAs 54  
Db 102 AGGGAGCGGGAGCGCAACCCCTGTGGAGAT-----GGAAACACTGA 143  
QY 54 pSerIleAsnSerAspAsp-----AlaAlaValAsnSerG1 66  
Db 144 GAGCGACAGAGCCCGGCGCTTGTGTACATGGACCCGACACATGGAGGCTGTGAAGCACA 203  
QY 66 nGlnGlnAsnGluLysGlnAspThrAspPheSerValLeuLysPheAlaTyIleArgProSe 86  
Db 204 ATTGCTGGAGCAAGCCAGGCAACCTGAGGAGCTGCTGGATCGGCCCATCGGGAGGC 263  
QY 86 rValProAlaHis-----Ar 91  
Db 264 TATACATCTTACCCATCACAAGACAAACCTCTGCCCCACCTCCCGAGGTTCTCTTGAG 323  
QY 91 gLysValLysGluSerProLeuSerSerAspThrIlePhe-----ArgGlnSer----- 107  
Db 324 CAGGACCCAGGAGCCATCCCTGGGAAACAGAAAGTTTTCATCATCCGCAAGTCCCTGTGT 383  
QY 108 -----HisAlaGlyLeuPheAsnLe 114  
Db 384 TGATGAGCTGATGGAGGTCCAGCATTTCCGACCATCTACCACATCTTCATCGCTGCCT 443  
QY 114 uCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsn-----LeuMe 132  
Db 444 GTGTGCTTTCATCATCAGCACCCCTGGCCATCGACTTCATTTGATGAGGCGAGGCTGTGTGT 503  
QY 132 tLysTyIleGlyTrpLeuLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTr 152  
Db 504 GGAGTTTGACTTACTATCTTCTGAGC---TTCGGACAGCTGCCATTCGCGCTGTGTGACCTG 560  
QY 152 p---ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleVa 171  
Db 561 GGTGCCCATGTTT-----CTGTCCACCCTGTGTGGCGCGTACCAGGCCCTA----- 606  
QY 171 lGluLysLeuAlaGlnArgLysCysIleProGluPro-----ValValVa 186  
Db 607 ----CGGCTGTGGCGCCAGGGGACCTGAGCAGCGAGCGCGCTGGGCTGGGTGTGTGTGT 662  
QY 186 lValLeuHisIleIleIleThrSerThrSerLeuPheTyIleProValLeuValIleLeuAr 206  
Db 663 AGCGGCCACGCGGT 713  
QY 206 gCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheSerCysValValTrpLe 226  
Db 714 GCATCAGCTCCCGCGCGCTCC-----CGTTGTGTCTGTGTGTCT 752  
QY 226 uLysLeuValSerTyIleAlaHisThrAsnTyIleAspMetArgAlaLeuThrLysLeuValG1 246  
Db 753 CGAGCAGGTTAGGTTCCTCTGATGAAAGTACTCTCTCTCTCTCTCTCTCTCTCTCTCT 795  
QY 246 uLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTyIleProTyIleAsnVal----- 263  
Db 796 -----GAGGCTGTGGGACCTCTGTGCCACAGGAGGTGAGGGGATCCAGGCCCGC 848  
QY 264 -SerPheLysSerLeuAlaTyIlePheLeuValAlaProThrLeuCysTyIleGlnProSerTy 283  
Db 849 CAGTTTCTCCAGCTACT 908  
QY 283 rProArgThrProTyIleArgLysGlyTrpLeu-----PheArgGlnLeuVally 300  
Db 909 CCTAGGACGCGCTATGTGAGGTGGAATATGTGTGGCAAGAACTTTTCCGACGCGCTGTGG 968  
QY 300 sLeuIleIlePheThrGlyValMetGlyPheIleIleAspGlnTyIleAsnProIleVa 320  
Db 969 ATGTGTGTCTATGCC-----TGCTTCATCTCTGGCGCGCTCTGTGTCTGTCTCT 1019



Db 2292 CT---CGGTTACTAATTCAGCTAG-CAGAAATCAAGCAGCTGTTCCTCAATACCTACAGTCA 2347  
QY 256 snMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProT 276  
Db 2348 AC-----CAGTATTGTACTCTTATTGCTCCTA 2377  
QY 276 hrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheA 296  
Db 2378 CCCCTATCTACCGTGACAGCTATCCAGGAATCCCACTGAAGATGGGTATGTGCTA 2437  
QY 296 rgGlnLeuValLysLeuIlelePheThrGlyValMetGlyPheIleleaspGlnTyrI 316  
Db 2438 TGAAGTTGGACAGCTCTGGTGGCTTTTCTATGTACTATCTTTGAAAGGCTTT 2497  
QY 316 leAsnProIleValGlnAsn---SerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaT 335  
Db 2498 GTGCCCTCTGTTTCGGAATATCAACAGGAGCCCTTCAGCGCTCGTCTGTGCTCTAT 2557  
QY 335 hrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyrC 355  
Db 2558 GT-----GTATTTAACTCCATCTGTCCAGGTGTGCTGATTTCTTCTTACTTTTGTG 2611  
QY 355 ysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluP 375  
Db 2612 CCTTTTGCAGCTGCTGCTCAATCCCTTGTGAGATGTACGCTTTTGGTGCACAGATG 2671  
QY 375 heTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetP 395  
Db 2672 TCTATAAGGATGTGTGAACTCCACGTCATACTCCAACCTATATAGAACCTGGAATCGG 2731  
QY 395 roValHisLysTrpMetIleArgHis-----LeuTyrPheProCysLeuA 410  
Db 2732 TGGTCCATGACTGCTATATATCTATCTTACAGGACTTTCTGCTGTTTCTTCCAGA 2791  
QY 410 rgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheH 430  
Db 2792 GATTC-----AAATCTGCTGCCATGTAGCTGCTTGTGATCTGCTGATGATAC 2842  
QY 430 lsGluLeuCysIleAlaValProCys-----HisIlePheLysLeuT 444  
Db 2843 ACGAATATGCTTGGCTGTT---TGCTTGAGCTTTTCTATCCGCTGCTGCTGCTCT 2899  
QY 444 rPalapheGlyIleMetPheGlnValProLeuValLeuIleThrAsnTyrLeuGlnA 464  
Db 2900 TCATGTTCTTTGGAATGGCTTTC-----AACTTCATTTGCTCA 2935  
QY 464 snLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSerIleLeuG 484  
Db 2936 ATGATAGTCGGAAGAACCCGATTTGGATGTTCTGATGTGGAC-----TTCTCTTTCTTG 2991  
QY 484 lyGlnProMetCysValLeuLeuTyrTyr 493  
Db 2992 GGCAATGGAGTCTTACTCTGCTTTTATTC 3020

## RESULT 14

US-08-509-187D-2  
; Sequence 2, Application US/08509187D  
; Patent No. 5834283  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Ta-Yuan and Chang, Catherine C.Y.  
; TITLE OF INVENTION: ACYL Coenzyme A: Cholesterol Acyltransferase  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/509,187D  
; FILING DATE: 31-JUL-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lamport Hammitte, Ann  
; REGISTRATION NUMBER: 34,858  
; REFERENCE/DOCKET NUMBER: DCI-033cpdv  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)742-4214  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4079 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-08-509-187D-2  
Alignment Scores:  
Pred. No.: 1.08e-32 Length: 4079  
Score: 376.00 Matches: 151  
Percent Similarity: 42.94% Conservative: 77  
Best Local Similarity: 28.44% Mismatches: 175  
Query Match: 14.03% Indels: 130  
DB: 26 Gaps: 26  
US-09-856-018B-16 (1-504) x US-08-509-187D-2 (1-4079)

QY 33 AsnSerProGlnThrThrAspSerSerGly-AspAspLeu-----AlaLys 48  
Db 1608 GATAGCAAGCAAGATAAAGTTGACAGAGGAGGAGGAATTGAAGCCATTTTATGAA 1667  
QY 48 saspSerGlySerAspAspSerIleAsnSerAspAlaAlaValAsnSerGlnGlnG 68  
Db 1668 GGAAGTGGGAGT-----CACTTTGATGATTTTGTGACCACATCTCATTTGAAA 1715  
QY 68 nAsnGluLysGlnAsp-----ThrAspPheSerValLeu-----79  
Db 1716 GTGAGCATCATTAATAATAGTGGTGGCTCTCACAACTTTTCTGTTCTGAAGCAGA 1775  
QY 80 -LysPheAlaTyrArgProSer-----ValPro-----88  
Db 1776 GAAAAACAACCATAGAGCGGAGGAGTTTGAGAGCACCTCCAGAACAAAGAAAGATTTTAT 1835  
QY 89 -AlaHisArgLysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerH 108  
Db 1836 TGCAAGCGGCTCTCTCTTAGATGAAGTCTTGAAGTGGACCATC-----1881  
QY 108 sAlaGlyLeuPheAsnLeuCysIleValValLeuValAla---ValAsnSerArgLeuI 127  
Db 1882 -AGAACAAATATATACATGTTTATGGCCCTCTCATCTCTTATCTCATGACACATGT 1940  
QY 127 eileGlu-AsnLeuMetLysTyrGlyTrpLeuIleLys-----139  
Db 1941 AGTAGATTACATTGATGAAGGAGGCTGGTGTGCTTACGCTCCCTGCTCTTATGCA 2000  
QY 140 -----SerGlyPheTrpPheSerSerLysSerLeuArgAspTrpProLeuPhe-----155  
Db 2001 TTTTGGCAAAATTC-----CTACCGTCTTTGGACCTG 2033  
QY 156 -----MetCysCysLeuSerLeuValValPheProPheAlaAlaIleValLysL 174  
Db 2034 GTGGATCATGTTCTGCTACATTTTTCAGTTCCCTATTTTCTGTTTCAACATTTGGCGCAC 2093  
QY 174 euAla---GlnArgLysCysIleProGluProValValValLeuHisIleIleIle 193  
Db 2094 TGGCTATAGCAAGAGTTCTCATCGCTGATCCGCTTCTCTCTCCATGGCTTTCTTTTCAT 2153



Db	1882	-AGAACATATATCATCATGTTTATTGGCCCTCCTCATCTCTTTATCTCTAGCACACTGTG	1940
Qy	127	elleGlu-AsnLeuMetLysTyrglyTrpLeulleLys-----	139
Db	1941	AGTATGATACATGATGAAGGAGCGCTGGTCTTCCAAAGTTACGCCCTCTCTTATGCA	2000
Qy	140	-----SerGlyPheTrpPheSerLysSerLeuArgAspTrpProLeuPhe-----	155
Db	2001	TTTGTGGCAAAATTC-----CTACCGTGTGTTGGACCTG	2033
Qy	156	-----NetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGluLysL	174
Db	2034	GTGGATCATCTCCGTCTACATATTTTCAGTCCCTATTTCTGTTTCAACATTTGGCGCAC	2093
Qy	174	euAla---GlnArgLysCysIleProGluProValValValValLeuHisIleIleIleI	193
Db	2094	TGCGTATACGAAGAGTTCATCCGCTCATCCGTCTCTCTCCATGGCTTCTTTTCAT	2153
Qy	193	hrSerThrSerLeuPheTyProValLeu-----	202
Db	2154	GATCTCCAGATTGGAGTTCATAGTCTTGGACCAACATATGTTGTTAGCATATCTCGC	2213
Qy	203	-----ValIleLeuArgCysAspSerAlaPheValSerGlyValThrL	217
Db	2214	CACCACTCCCGGTTCATCATTTATGAGCAGAT---CGTTTTGTA-----	2257
Qy	217	euMetLeuPheSerCysValValTrpLeuLysLeuValSerTyAlaHisThrAsnTyra	237
Db	2258	-----ATGAGGCCCACTCATTTGTCAGAGAACGCTGC	2291
Qy	237	spMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeu---LeuAspThrLeuA	256
Db	2292	CT---CGGCTACTAATTCAGCTAAG-GAGAAATCAAGCACGTGTCCAATACCTACAGTCA	2347
Qy	256	snMetAspTyrrProTyrrAsnValSerPheLysSerLeuAlaTyrrPheIleuValAlaProT	276
Db	2348	AC-----CAGTATTTGTACTCTTTATTTGCTCCTA	2377
Qy	276	hrLeuCysTyrrGlnProSerTyrrProArgThrProTyrrIleArgLysGlyTrpLeuPheA	296
Db	2378	CCCTTATCTACCGTGACAGCTATCCAGGAATCCCATGTAAGATGGGGTATATGTTGCTA	2437
Qy	296	rgGlnLeuValLysLeulleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrrI	316
Db	2438	TGAAGTGTGCACAGGCTTTGGTGTCTTTTCTATGTGTACTACATCTTTGAAAGGCTTT	2497
Qy	316	leAsnProIleValGlnAsn---SerGlnHisProLeuLysGlyAsnLeuLeuTyrrAlaT	335
Db	2498	GTCCCCCTGTGTTTCGGAAATATCAACAGAGAGCCCTTCAGCGCTGTTGTTCTGGTCTAT	2557
Qy	335	hrGluArgValLeuLysLeuSerValProAsnLeuTyrrValTrpLeuCysMetPheTyrrC	355
Db	2558	GT-----GTATTTAACTCCATCTTGCCAGGTGTGCTGATTCTCTCTTACTTTTITTTG	2611
Qy	355	ysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluP	375
Db	2612	CCTTTTTGCATCGTGGCTCAATGCCTTTGCTGAGATGTTACGCTTTGGTGACAGAGATG	2671
Qy	375	heTyrrLysAspTrprPrAsnAlaLysThrValGluAspTyrrTyrrArgMetTrpAsnMetP	395
Db	2672	TCATTAAGGATTGGTGAACCTCCAGTCATATCCCACTATTATAGAACCTCGGAATGTGG	2731
Qy	395	roValHisLysTrpMetIleArgHis-----LeuTyrrPheProCysLeuA	410
Db	2732	TGGTCCATGACTGGCTATATTTACTATGCTTACAGGACTTTTCTCTGGTTTTCTCCAAGA	2791
Qy	410	rgHisGlyLeuProLysAlaAlaLeuLeuLeulleIleAlaPheValSerAlaLeuPheH	430
Db	2792	GATTC-----AAATCTGCTGCCATGTGTAGCTGCTTGTGTATCTGCTGTAGTAC	2842
Qy	430	isGluLeuCysIleAlaValProCys-----HisIlePheLysLeuT	444
Db	2843	ACGAATATGCTTGGCTGTT---TGCTTGAGCTTTTCTATCCCGTCTGTCTGTCTGCTCT	2899

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 21, 2003, 19:43:31 ; Search time 1672 Seconds  
(without alignments)  
4881.893 Million cell updates/sec

Title: US-09-856-018b-16  
Perfect score: 2680  
Sequence: 1 MAISDEPESVATLHNSLR.....QPMCVLLYHDLNRKGLD 504

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xl  
-Q=/cgn2\_1/spool/US09856018/runat\_20022003\_091206\_8351/app\_query.fasta\_1.647  
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcg -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09856018 @CGN\_1\_1\_2874 @runat\_20022003\_091206\_8351 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_vrt:\*  
23: em\_gss\_fun:\*  
24: em\_gss\_mam:\*  
25: em\_gss\_mus:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1321	49.3	1572	11	AY105372	Zea mays
C	1187	44.3	961	14	BQ510367	AY105372 Zea mays
3	1107.5	41.3	1181	11	AY110660	AY110660 Zea mays
4	1038.5	38.8	729	13	BI422326	BI422326 EST332932
C	952	35.3	586	14	BQ148998	BQ148998 NF086D09F
6	885	33.0	855	10	AW349274	AW349274 GM210004B
7	880	32.8	572	13	BQ309699	BM309699 sak65f03.
8	871	32.5	741	14	BQ989822	BQ989822 OGG20J23.
C	854	31.9	729	13	BQ322102	BQ322102 BJ322102
10	823.5	30.7	572	14	BQ124305	BQ124305 EST609881
11	800	29.9	606	13	BQ316561	BQ316561 BJ316561
12	784	29.3	539	13	BI422212	BI422212 EST532878
13	761	28.4	540	9	AJ470192	AJ470192 AJ470192
14	732.5	27.3	561	10	AW035727	AW035727 EST281881
15	729	27.2	433	9	AI441040	AI441040 sa58f02.y
16	707	26.4	626	10	AV926912	AV926912 AV926912
17	691	25.8	862	12	BG321213	BG321213 Zm04_0590
18	660	24.6	682	10	BE247899	BE247899 NF038D11D
19	657	24.5	577	10	AW586836	AW586836 EST318459
20	652	24.3	763	13	BI308446	BI308446 EST529856
21	643	24.0	655	14	BQ988311	BQ988311 OGG19C15.
22	635	23.7	480	9	AJ470191	AJ470191 AJ470191
23	632	23.6	480	9	AJ470190	AJ470190 AJ470190
24	626.5	23.4	459	14	BQ123670	BQ123670 EST609246
25	620	23.1	634	12	BF634363	BF634363 NF059D06D
26	608	22.7	1043	14	BM919422	BM919422 AGENCOURT
27	601.5	22.4	673	14	BQ985554	BQ985554 OGG19N20.
28	597.5	22.3	1064	13	BM476315	BM476315 AGENCOURT
29	597	22.3	508	14	BU009402	BU009402 QGJ10F20.
30	579	21.6	555	13	BJ268713	BJ268713 BJ268713
31	572	21.3	685	14	BQ861203	BQ861203 QCCL17N02.
32	568.5	21.2	639	9	AA042298	AA042298 24635 CD4
33	559.5	20.9	619	14	BM739263	BM739263 K-EST00008
34	559	20.9	515	10	AV925760	AV925760 AV925760
35	558.5	20.8	968	14	BQ670343	BQ670343 AGENCOURT
36	555.5	20.7	627	10	AV956927	AV956927 AV956927
37	549	20.5	656	14	BQ862755	BQ862755 QCC21P16.
38	543	20.3	819	13	BI225365	BI225365 602950108
39	540.5	20.2	661	14	BQ042483	BQ042483 UI-M-EMO.
40	539	20.1	655	10	AW775077	AW775077 EST334228
41	529.5	19.8	585	12	BQ078967	BQ078967 H3034B09-
42	524	19.6	574	12	BF003479	BF003479 EST431977
43	516.5	19.3	570	14	BM829896	BM829896 K-EST0103
C	516	19.3	617	10	AV930897	AV930897 AV930897
C	501	18.7	857	9	AL520752	AL520752 AL520752

ALIGNMENTS

RESULT	1	AY105372	PC0148220	mRNA	linear	HTC	25-MAY-2002
LOCUS	AY105372	Zea mays	1572 bp	mRNA	linear	HTC	25-MAY-2002
DEFINITION	AY105372	Zea mays	PC0148220	mRNA	sequence.		
ACCESSION	AY105372	Zea mays	PC0148220	mRNA	sequence.		
VERSION	AY105372.1	GI:21208450					
KEYWORDS	HTC						
SOURCE	Zea mays.						
ORGANISM	Zea mays						
REFERENCE	1	(bases 1 to 1572)					
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.						

TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of  
Overgo Probes

JOURNAL Unpublished (2002)

REFERENCE 2 (bases 1 to 1572)

AUTHORS Coe,E.C.

TITLE Direct Submission

JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of

Missouri, Columbia, MO 65211, USA

FEATURES Location/Qualifiers

1..1572

/organism="Zea mays"

/db\_xref="MaizeDB:638748"

/db\_xref="taxon:4577"

/clone="PCOI48220"

/clone\_lib="Maize Mapping Project/DuPont Consensus

Library"

/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed Dupont contigs; this resource was  
assembled by Dupont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"

BASE COUNT 438 a 294 c 285 g 555 t

ORIGIN

Alignment Scores:

Pred. No.: 3.79e-133 Length: 1572

Score: .1321.00 Matches: 235

Percent Similarity: 66.90% Conservative: 54

Best Local Similarity: 54.40% Mismatches: 72

Query Match: 49.29% Indels: 71

DB: 11 Gaps: 1

US-09-856-018b-16 (1-504) x AY105372 (1-1572)

Qy 142 PheTppPheSerLysSerLeuArgAspTrpProLeuPheMetCysCysLeuSerLeu 161  
Db 1 TTTTGGTTAATGCTACATCATGTCGAGACTGGCCACTGCTAATGTGTGCTTAGTCTA 60  
Qy 162 ValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIlePro 181  
Db 61 CCCATATTTCCCTTGGTGCACTTCCAGTCGAAAGTTGGCATTCACAAATCTCGTTAGT 120  
Qy 182 GluProValValValLeuHisIleIleIleIleIleIleIleIleIleIleIleIleIle 201  
Db 121 GATCGTCTACTACCTGTTTTCACATCTCTTTTACACATTTTGAATGTATATCCAGTG 180  
Qy 202 LeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheSer 221  
Db 181 CTCGGAATCTTAAGTGTGATCTCCAGTTTATCAGGCTTTGTGTTGATGTTTATTCGCC 240  
Qy 222 CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAlaLeu 241  
Db 241 TGCATTTGTTGGCTGAAGCTGTATCTTTTGCACATACAAACCATGATATAGAAACATG 300  
Qy 242 ThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTyrProTyr 261  
Db 301 ATCACAAGCGCAGAGAGTTGATTAATGAAGTACCGCGGCTGGCATAGATAATTTACAA 360  
Qy 262 AsnValSerPheLysSerLeuAlaThrPheLeuValAlaProThrLeuCysTyrGlnPro 281  
Db 361 GCTCAACTCTGGGAGTCTACATACATTCATGATGGCTCCGACACTCTGTTATCAGCCA 420  
Qy 282 SerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeuLysLeu 301  
Db 421 AGTTATCTCGCAACACCTTATGTTAGAAAGTTGGCTGGCTCCGTCAGTATTTCTATAC 480  
Qy 302 IleIlePheThrGlyValMetGlyPheIleIleIleIleIleIleIleIleIleIleIleIle 321  
Db 481 TTGATATTTACTGCTCTCAAGATTCATTTATGAGCAATACATAAATCTTATTTGTTG 540  
Qy 322 AsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLeu 341

Db 541 AACTCTCAACATCCATTGATGGGAGGATTACTGAATGCTGTAGAGACTGTTTTGAAGCTC 600  
Qy 342 SerValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeu 361  
Db 601 TCATTACCAAAATGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560  
Qy 362 AsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrpTrpAsn 381  
Db 661 AACATACTGCTGAGATCTTCGATTGCTGAGCAGGAGATTTCTACAAAGACTGGTGAAT 720  
Qy 382 AlaLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTrpMetIle 401  
Db 721 GCAAGACAAATGATGAGTACGAGAAATGAGCATGCCCTGTCATAAATGGATTGTT 780  
Qy 402 ArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAlaLeuLeuIle 421  
Db 781 CGTCATATATTTTCTGTCATGCGAATGATATATCAAGGAAGTGTCTGTTTTTATA 840  
Qy 422 AlaPheLeuValSerAlaLeu-Phe- 429  
Db 841 TCGTTCTTTGTTCTGCTACTTCTATGAGGTAACCTATTATTCTTTTCACTCTTCATCT 900  
Qy 429 ----- 429  
Db 901 GCATATATTAATATATATAGTCTCTATTTCAAATGTGTCTTTCGAGTTTCGACATGCT 960  
Qy 429 ----- 429  
Db 961 TTTGTTCAAACTTACCAGCTGTAGATTACTTGGATGAAGTGTCTATATATAAATCAATA 1020  
Qy 430 -----HisI 431  
Db 1021 TTTCACAATCCAGTCCCTTTCGAGAAATATATATACATTTTGTGCTTTGTACACCA 1080  
Qy 431 uLeuCysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPh 451  
Db 1081 GTATGCGTTCGAGTTCCCTGCCACATATCAAGTCTCTGGGCTTTCTTAGGAATCATGCT 1140  
Qy 451 eGlnValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetVa 471  
Db 1141 TCAGATTTCCCTCATCATATTGACATCATCTCAAAATAAATTCAGTGACACAAATGGT 1200  
Qy 471 IglyAsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnPrometCysValLeuLe 491  
Db 1201 TGGCAATATGATCTTTTGTGTTTTTCTGATATACGGCAGCAATGTGTGTCTATT 1260  
Qy 491 uTyrTrpHisAspLeuMetAsnArgLysGlyLys 502  
Db 1261 GTATTACCATGATGTGATGAACCGGACTGAGAAAG 1294  
RESULT 2  
BQ510367/c  
LOCUS BQ510367.2  
DEFINITION EST617782 Generation of a set of potato cDNA clones for microarray  
analyses mixed potato tissues Solanum tuberosum cDNA clone STMHJ76  
3' end, mRNA sequence.  
ACCESSION BQ510367  
VERSION BQ510367.2 GI:21926064  
KEYWORDS EST.  
SOURCE potato.  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asterales; euasterids I; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 961)  
Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,  
Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and  
Karamycheva,S.A.  
Generation of a set of potato cDNA clones for microarray analyses  
Unpublished (2002)  
On Jun 10, 2002 this sequence version replaced gi:21369236.  
Other ESTs: EST617781  
Contact: Robin Buell



The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Email: potato@tigr.org  
 This clone is available through the Research Genetics, contact the  
 Research Genetics for further information 1-800-711-6195 or  
 cdna@resgen.com

Seq primer: T7.

Location/Qualifiers

1. 961

/organism="Solanum tuberosum"

/cultivar="Kennebec or Binjete"

/db\_xref="taxon:4113"

/clone="STMHJ76"

/clone\_lib="Generation of a set of potato cDNA clones for

microarray analyses mixed potato tissues"

/tissue\_type="mixed tissues"

/lab\_host="SOLR"

/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:

XhoI; supplier: Combination of untreated and Phytophthora  
 infestans-treated libraries of stolons, leaves, leaflets,  
 axillary buds of stem explants, petioles, germinating eyes  
 tubers, or roots."

324 a 205 c 181 g 251 t

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 7,61e-119 Length: 961  
 Score: 1187.00 Matches: 205  
 Percent Similarity: 88.24% Conservative: 35  
 Best Local Similarity: 75.37% Mismatches: 28  
 Query Match: 44.29% Indels: 4  
 Db: 14 Gaps: 1

US-09-856-018B-16 (1-504) x BQ510367 (1-961)

Qy 230 SerTyrAlaHisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGlu 249  
 Db 961 TCTATGACATACAAATATGATGATGAGACAGCTTGCAGAGTCTGTGATGAGGGTGAA 902  
 Qy 250 AlaLeuLeuAspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAla 269  
 Db 901 -----AATCCGAATCAACTACTCTTCAATGTTAGTTTCAAGAGTTTGCT 854  
 Qy 270 TyrPheLeuValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIle 289  
 Db 853 TACTTCATGTTGCTCCAACTTATGCTATGCTAGCTTAGCTATCCCTCGCTCGATCCCAT 794  
 Qy 290 ArgLysGlyTyrPheArgGlnLeuValLysLeuIlePheThrClyValMetGly 309  
 Db 793 CGAAGGTTGCTGCGCCCACTCACTCAAGCTGGTAATTTTACAGGATTAAATGGGA 734  
 Qy 310 PheIleLeuAspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGly 329  
 Db 733 TTTATCATTGACGAGTATATTAAACCGATTGTGCGAAGCTCACACATCCATTCGAAGA 674  
 Qy 330 AsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrp 349  
 Db 673 AACCTTTTATACGCCATCGAGAGGGTATTGAAGCTTTTCAGTTTCCAAATTTATATGCTGG 614  
 Qy 350 LeuCysMetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArg 369  
 Db 613 CTCGTCATGTTCTACAGGCTCTTTCATCTTGGCTAAATATACTTGCAGAAATTCGCGGA 554  
 Qy 370 PheGlyAspArgGluPheTyrLysAspTrpAsnAlaLysThrValGluAspTyrTrp 389  
 Db 553 TTTGGGATCGTGAAGTCTACAAAGTTGGTGAATGCAAAACAAATTCATGATGATTGG 494  
 Qy 390 ArgMetTrpAsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeu 409  
 Db 493 AGACTTTGGAATATGCCGTGATCAATAGTGGATGGTTCGCCCATTTATTTCCATGCTTA 434  
 Qy 410 ArgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPhe 429

Db 433 AGGAATGCATACCTAAGGAGTTCGAATGGTGATCTCTTTCTTTATATCTGCTCTTTTC 374  
 Qy 430 HisGluLeuCysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIle 449  
 Db 373 CATGAGCATGATGCTGTTCTTCCCTGCGCTATTCAAGTTTGGCATTCCTTGGATC 314  
 Qy 450 MetPheGlnValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSer 469  
 Db 313 ATGTTTTCAGATTCCCTTGGTCATACCTACGAACTTCTCCAAACAAAGTTCAAAACTCG 254  
 Qy 470 MetValGlyAsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysVal 489  
 Db 253 AATGGGCAACATGACATTCGGTGTCTTCTGCATTTGGTCAACCAATGTTGTGG 194  
 Qy 490 LeuLeuTyrTyrHisAspLeuMetAsnArgLysGly 501  
 Db 193 CTCCTGTATTACCATGATGATGATGAATAGAAATGCT 158  
 RESULT 3  
 AY110660 1181 bp mRNA linear HTC 26-MAY-2002  
 DEFINITION Zea mays CL739\_1 mRNA sequence.  
 ACCESSION AY110660  
 VERSION AY110660.1 GI:21215250  
 KEYWORDS HTC.  
 SOURCE Zea mays.  
 ORGANISM Zea mays.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
 clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 1181)  
 AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,  
 Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
 TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of  
 Overgo Probes  
 JOURNAL Unpublished (2002)  
 REFERENCE 2 (bases 1 to 1181)  
 AUTHORS Coe,E.C.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA  
 FEATURES  
 source  
 1..1181  
 /organism="Zea mays"  
 /db\_xref="MaizeDB:632842"  
 /db\_xref="taxon:4577"  
 /clone="CL739\_1"  
 /clone\_lib="Maize Mapping Project/DuPont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 contigs to seed DuPont contigs; this resource was  
 assembled by DuPont as part of a collaboration for the  
 overgo addressing of BACS in conjunction with the Maize  
 Mapping Project"  
 BASE COUNT 322 a 240 c 275 g 327 t 17 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 5,28e-110 Length: 1181  
 Score: 1107.50 Matches: 197  
 Percent Similarity: 83.39% Conservative: 34  
 Best Local Similarity: 71.12% Mismatches: 39  
 Query Match: 41.32% Indels: 7  
 Db: 11 Gaps: 3  
 US-09-856-018B-16 (1-504) x AY110660 (1-1181)  
 Qy 229 ValSerTyrAlaHisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGly 248  
 Db 12 GTCTCTTATGCACATACAAATATCATATAGGGTATTCTCCAAAGTACTGAGAGGGT 71  
 Qy 249 GluAla-----LeuLeuAspThrLeuAsnMetAspTyrProTyrAsnValSerPhe 265



```

|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
590 GCATTCCTTGGATCATGTTTCAGATTCCTGTCATACCTAAGCACTTCCTGCAAAAC 649
QY 465 LysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSerIleLeuGly 484
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
650 AAGTTCAAAACCTGAATGTTGGCAACATGACATCTCTGTTCTTCTTCTGCTATGTTGGT 709
QY 485 GlnProMetCysValLeuLeu 491
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
710 CAACCATG-TGTGTGCTTCG 729

RESULT 5
BQ148998
LOCUS
DEFINITION
NF086D09FLF1077 Developing flower Medicago truncatula cDNA clone
NF086D09FL 5', mRNA sequence.
ACCESSION
VERSION
BQ148998.1 GI:20286057
KEYWORDS
SOURCE
barrel medic.
ORGANISM
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 586)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula flower library
Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert length: 586 Std Error: 0.00
Plate: 086 row: D column: 09
Seq primer: TCACACGAGAAACAGCTATGAC.
FEATURES
Location/Qualifiers
1..586
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF086D09FL"
/clone_lib="Developing flower"
/tissue_type="Developing flower"
/dev_stage="Developmentally pooled. Contains a mixture of
very young, developing, fully-opened flowers and flowers
in early transition into pods."
/note="Vector: Lambda Zap; cDNA was prepared from polyA+
enriched, pooled samples of equivalent amounts of total
RNA from very young, developing, fully-opened flowers and
flowers transitioning into pods. The cDNA was
directionally ligated into the Uni-Zap XR vector
(Stratagene) and packaged using the Gigapack III Gold
packaging extracts. Phagemids containing cDNA inserts were
in vivo excised from the recombinant Uni-Zap XR vector
using ExAssist helper phage and the E. coli strain
XL1-Blue MRF' (Stratagene). Excised plasmids were plated
using SOLR cells."
BASE COUNT 134 a 115 c 132 g 202 t 3 others
ORIGIN

Alignment Scores:
Pred. No.: 1 69e-93 Length: 586
Score: 952.00 Matches: 165
Percent Similarity: 93.81% Conservative: 17
Best Local Similarity: 85.05% Mismatches: 12
Query Match: 35.52% Indels: 0
DB: 14 Gaps: 0

```

```

US-09-856-018B-16 (1-504) x BQ148998 (1-586)
QY 271 PheLeuValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArg 290
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
3 TTCATGTTGCTCTCATATTATGTCACGCAAGCTATCTCGACACCTTCGCTCGA 62
QY 291 LysGlyTrpPheArgGlnLeuValLysLeuIlePheThrGlyValMetGlyPhe 310
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
63 AAGGTTGGGTCTGTCACAACTTCTCAAGCTGCTCATATTACAGAGTATTGCGATT 122
QY 311 IleIleAspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsn 330
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
123 ATAATAACAATATATATGATCTTATGTCAGAAATTCACACATCATCGAAGGAAAC 182
QY 331 LeuLeuTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeu 350
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
183 CTTCTATATGCCATGAGAGAGTTCTGAAGCTTCTGTCCTCAAAATGTTATGTCGCTG 242
QY 351 CysMetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPhe 370
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
243 TGCATGTTCTATGCTTTTCCATCTTTGGTTAAATATACTTGGGAGCTTCTCCGTTT 302
QY 371 GlyAspArgGluPheTyrLysAspTyrTrpAsnAlaLysThrValGluAspTyrTrpArg 390
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
303 GGTGATCGTGAGTCTACAAGATGCTGGATGCCCAACGGTGAAGAGTATGGAGG 362
QY 391 MetTrpAsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArg 410
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
363 ATGTGGAATATGCTGTGCACAAATGATGCTGTCACGTGTTATTTCCCTGCATAAGG 422
QY 411 HisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHis 430
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
423 TTTGGTATACCAAGGCTGCTGCTTGTGACTCTTCTCTGTTCTGCTGTTCCAT 482
QY 431 GluLeuCysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyMet 450
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
483 GAGTTATGCAATGCTGTCTCTGCCCATGTTCAAGTGTGGGCTTTTATTGGAATTATG 542
QY 451 PheGlnValProLeuValLeuIleThrAsnTyrLeuGlnAsn 464
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
543 TTCCANGTCTTTGGTCTTGATCACCAATTACCTGAAAAAT 584

RESULT 6
AW349274/c
LOCUS
DEFINITION
GM210004B21H12 Gm-r1021 Glycine max cDNA clone linear EST 04-OCT-2000
mRNA sequence.
ACCESSION
VERSION
AW349274.1 GI:6846984
KEYWORDS
SOURCE
soybean.
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
1 (bases 1 to 855)
Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Corryell, V.,
Erpelting, J., Raph, C., Shoop, E., Pardini, J., Liu, L., and Lewin, H.
A Functional Genomics Program for Soybean (NSF 9872565)
Unpublished (1999)
Other_ESTs: A1441040
Contact: Vodkin, L.O., PI, A Functional Genomics Program for
Soybean (NSF 9872565)
Lewin, H. A., Director, Keck Center for Comparative and Functional
Genomics
University of Illinois
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
Tel: (217) 244-6147
Fax: (217) 333-4582
Email: l-vodkin@uiuc.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134. For further information

```



```

Pred. No.: 1.22e-85 Length: 572
Score: 880.00 Matches: 179
Percent Similarity: 93.40% Conservatives: 5
Best Local Similarity: 90.86% Mismatches: 6
Query Match: 32.84% Indels: 7
DB: 13 Gaps: 3

US-09-856-018B-16 (1-504) x BM309699 (1-572)

QY 5 AspGluProGlu-SerValAlaThrAlaLeuAsnHisSerSerLeuArgArgProSe 24
Db 2 GATGAGCCTGATTCGTACCCACTGCTCAACCACTCTCCCTGCGCGCGTCCAC 61
QY 24 rAlaThrSerThrAlaGlyLeuPheAsnSerProGluThrThrThrAspSerGlyAs 44
Db 62 CGCC-----GCTGGCCCTCTCAATTCGCCGAGACGACACCCAGCTTCCGGTGA 112
QY 44 pAspLeuAlaLysAspSerGlySerAspSerIleAsnSerAspAlaAlaValAs 64
Db 113 TGACTTGGCCAGAGATCCCGTTCCGAGACTCCATCAGCAGCGAC-----GCCGCCAA 166
QY 64 nSerGlnGlnAsnGluLysGlnAspThrAspPheSerValLeuLysPheAlaTyAr 84
Db 167 TTCGAACCGCAA--CAAAAACAAGACACTGATTTCTCGCTCCTCAAAATTCGCTACCG 223
QY 84 gProSerValProAlaHisArgLysValLysGluSerProLeuSerSerAspThrIlePh 104
Db 224 TCCTTCGCTCCCGCTCATCGCAAGTGAAGGAAAGTCCGCTCAGCTCCGACACCACTTT 283
QY 104 eArgGlnSerHisAlaGlyLeuPheAsnLeuCysIleValValValAlaValAsnSe 124
Db 284 CCCTCAGAGTCACGCGGCCCTCTCAACCTCTGTATAGTAGTCTTGTCTGTGAATAG 343
QY 124 rArgLeuIleGluAsnLeuMetLysTyGlyTrpLeuIleLysSerGlyPheTrpPh 144
Db 344 CCGACTCATCATTCAGCAATTAATGAAGTATGTTGCTTGCATCAAACTGCTTTGTT 403
QY 144 eSerSerLysLeuArgAspTrpProLeuPheMetCysLeuSerSerLeuValValPh 164
Db 404 TAGCTCAAGTCATTCAGAGACTGGCCCCCTTCATGCTGTGTCTTCTCTTGTGTATT 463
QY 164 eProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIleProGluProVa 184
Db 464 TCCTTTGTGCTATATATAGTGGAGAAGTGGACAGCAGAGTATACCCGAGACAGN 523
QY 184 lValValValLeuHisIleIleIleThrSerThrSerLeuPheTyPro 200
Db 524 TGTGTGGACTTCATATAATCATTAATACCTCAGCTTCACTTTCTATCCA 572

RESULT 8
BQ998922
LOCUS BQ998922
DEFINITION BQ998922 741 bp mRNA linear EST 22-AUG-2002
            OGG20J23.yg.ab1 OG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
            OGG20J23, mRNA sequence.
ACCESSION BQ998922
VERSION BQ998922.1 GI:22433318
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
            Lactuca.
REFERENCE 1 (bases 1 to 741)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
        Lin,H., van Damme,M., Lavelle,D., Chevaller,P., Ziegler,J., Ellison
        P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
        Church,S., Jackson,L. and Bradford,K.
        Lettuce and Sunflower ESTs from the Compositae Genome Project
        http://comgenomics.ucdavis.edu/
        Unpublished (2002)
        Contact: Alexander Kozik [R.W.Michelmore]
        Department of Vegetable Crops, R.W.Michelmore Lab

```

```

FEATURES
            Location/Qualifiers
            source
            1..741
            /organism="Lactuca sativa"
            /cultivar="L.serriola"
            /db_xref="taxon:4236"
            /clone_lib="OGG20J23"
            /lab_host="E.coli"
            /note="Vector: pBRCDNASFIAB: The library was constructed
            from 10 different sources of RNA from a single genotype.
            Separate cDNAs were generated using primers that
            incorporated unique 5' and 3' tags to distinguish each
            source of RNA. cDNAs were then pooled, size-fractionated,
            directionally cloned into a custom medium-copy vector and
            transformations made with four size classes to minimize
            size bias. Details of each source of RNA and library
            construction can be obtained at http://cgpdb.ucdavis.edu/
            TAG_LIB=OG_EFGHJ lettuce serriola
            TAG_TISSUE=flowers pre-fertilized
            TAG_SEQ=GGTTCACGGG"
BASE COUNT 197 a 145 c 144 g 255 t
ORIGIN
Alignment Scores:
Pred. No.: 1.79e-84 Length: 741
Score: 871.00 Matches: 167
Percent Similarity: 81.56% Conservatives: 32
Best Local Similarity: 68.44% Mismatches: 45
Query Match: 32.50% Indels: 1
DB: 14 Gaps: 0

US-09-856-018B-16 (1-504) x BQ998922 (1-741)
QY 77 SerValLeuLysPheAlaTyArgProSerValProAlaHisArgLysValLysGluSer 96
Db 10 AGTGTCGACATATGCGTATCGACCGCTCTTCCAGCTCATCGACATTTAAAGAGTCT 69
QY 97 ProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPheAsnLeuCysIle 116
Db 70 CCTTAAGTCTGAGCGCCATTTTCAAGCAGAGTCATCGAGACTCTTTAACCTTTGCATA 129
QY 117 ValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMetLysTyGlyTrp 136
Db 130 GTGGTCTTAGTTCAGTCAATGGAGCAGCTCATCATTTGAGAATCTGATGAAGTATGATTA 189
QY 137 LeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrpProLeuPheMet 156
Db 190 TTGATAAATCCCAATTTTGGTTCAGTTCAGATCATTCAGAGATTCGCCACTCTTTATG 249
QY 157 CysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGluLysLeuAlaGln 176
Db 250 TGTCGCTCACTCTCTCCAACTTCCCGCTACTCTACTATATATGTTGAAAATTAGCATGG 309
QY 177 ArgLysCysIleProGluProValValValValLeuHisIleIleIleThrSerThrSer 196
Db 310 CAAAACGATATTTTCAGACCCCTGTTGTAATCACTCTCCATATTCCTAACCAACTACAACGTCA 369
QY 197 LeuPheTrpProValLeuValIleLeuArgCysAspSerAlaPheValSerGlyValThr 216
Db 370 ATCTGATCCAGTTTCATGATTTTCAGGTTTCAGGTTTCAGTTCAGTTCATGAGCGGTATCA 429
QY 217 LeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyAlaHisThrAspTy 236
Db 430 TTGATGCTTGTGCTTGCATTAATTTGGTTGAAGTTGGTTCTTTCTTTTGTGCATGCAATAT 489

```

University of California at Davis (UCD)  
 Asmundson Hall, UCD, Davis, CA 95616, USA  
 Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]  
 belongs to contig OG\_CA\_Contig2177, see http://cgpdb.ucdavis.edu/  
 for details.  
 Plate: OGG20 row: J column: 23.



Qy	325	HisProLeuLysGlyAsnLeuLeuLysrAlaThrGluArgValLeuLysLeuSerValPro	344
		:	
Db	480	CATCCTTTTAAGGGTAACATTTCTATATGCCATTGAAAGAGTCTGTGAAGCTGCTGTGTTCCA	539
Qy	345	AsnLeuLysrValTrpLeuCysMetPheTyrCys	355
Db	540	AATCTATATGTGTGGCTCTGCATGCTTCTACTGC	572
RESULT	11		
LOCUS	BJ316561	606 bp	mRNA linear EST 09-APR-2002
DEFINITION	BJ316561 Y. Ogihara unpublished cDNA library, Wh_yf Triticum aestivum cDNA clone whyf23117 5', mRNA sequence.		
ACCESSION	BJ316561		
VERSION	BJ316561.1	GI:20121748	
KEYWORDS	EST.		
SOURCE	bread wheat.		
ORGANISM	Triticum aestivum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Triticum.		
AUTHORS	Ogihara,Y. and Murai,K.		
TITLE	Expressed genes in Triticum aestivum		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: Tadasu Shin-i Center for Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel. 81-559-81-6856		

```

FEATURES
    source
        1. .606
            /organism="Triticum aestivum"
            /cultivar="Chinese Spring"
            /db_xref="taxon:4565"
            /clone="chvf3117"

```

```

/Clone_lib=y. Ogihara unpublished cDNA library, Wh_yf#
/tissue_type="spikelet at early flowering"
/dev_stage="Feekees' scale 6"
/notes=Vector: Lambda Uni-ZAP XR, excised phagemid;
Site.1: EcoRI; Site.2: XhoI; Plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akhunov
in J. Bvorak Lab). Total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give pBluescript phagemids in the TJ Close Lab
at the University of California, Riverside (Akhunov, Chin
Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors).
BASE COUNT      170 a      113 c      125 g      198 t
ORIGIN
Alignment Scores:
Pred No.:      7 5e-77      Length:      606
Score:      800.00      Matches:      137
Percent Similarity:      81.46%      Conservative:      30
Best Local Similarity:      66.83%      Mismatches:      32
Query Match:      29.85%      Indels:      6
DB:      13      Gaps:      2
US-09-856-018B-16 (1-504) x BJ316561 (1-606)
Qy 214 GlyValThrLeuMetLeuPheSerCysValValrPrLeuLysLeuValSerTyrAlaHis 233
||| ||||| :||:|||||:|||||:|||||:|||||:|||||:
Db 1 GGTGTTTGGTGTGTTTATTGTCGATTTGGTGCATTCGATCTCTTTTGGCCAT 60
Ov 234 ThrAsnTyrAspMetArgAlaLeuThr-----TvsLeuValGluTvsGluVala 250

```







AUTHORS Saren, A.-M., Tanskanen, J., Paullin, L. and Schulman, A.H.

TITLE Barley EST's

JOURNAL Unpublished (2002)

COMMENT Contact: Schulman AH

Institute of Biotechnology

University of Helsinki

P.O.Box 56 (Vilinkaari 6A), University of Helsinki FIN-00014,

Finland.

Location/Qualifiers

1. 540

/organism="Hordeum vulgare"

/db\_xref="taxon:4513"

/clone="S000800182D08F1"

/clone\_lib="S00008"

/tissue\_type="Callus"

/note="Callus K19"

BASE COUNT 149 a 102 c 118 g 171 t

ORIGIN

Alignment Scores:

Pred. NO.: 1.14e-72 Length: 540

Score: 761.00 Matches: 128

Percent Similarity: 87.15% Conservative: 28

Best Local Similarity: 71.51% Mismatches: 23

Query Match: 28.40% Indels: 0

DB: 9 Gaps: 0

US-09-856-018B-16 (1-504) x AJ470192 (1-540)

Qy 230 SerTyrAlaHisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGlu 249

Db 2 TCTTTTGTCTATACAAATCATGATATAGGTATTTGCCAAAGATTTGAAAGGGTCT 61

Qy 250 AlaLeuLeuAspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAla 269

Db 62 ACACATGGCAGTTCCTATCGATCAAGAAACCATTAAGGTCCAACTACCAACAGTGTGTG 121

Qy 270 TyrPheLeuValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIle 289

Db 122 TATTTCATGTTGGGCCCAACACATTTGTACCAGCCAAAGTTATCCCGACAGCATTTGTT 181

Qy 290 ArgLysGlyTrpLeuPheArgGlnLeuValLysLeuIlePheThrGlyValMetGly 309

Db 182 AGGAAGGCTGGGTGGCCACGACGCTATATAAATGCATAGTTTACAGGCTGTATGGC 241

Qy 310 PheLeuLeuAspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGly 329

Db 242 TTCATAATTGACCAATACATTAATCCAATTGTGCAGAAATTCAGCATCCATTGAAAGGA 301

Qy 330 AsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrp 349

Db 302 AATTTCCTGGATGCTATTGAGAGAGTCTGAACTCTCAGTGCCGACATTTGTATGTGG 361

Qy 350 LeuCysMetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArg 369

Db 362 CTTTGATGTTCTATTGCTTTTCCATCTGTTGATATATCTGCCCACTCCCTCGT 421

Qy 370 PheGlyAspArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrp 389

Db 422 TTTGGTGATCGTAATTCATAAGGACGTGGTGAATGCCAGAACAGTTGAAGAGTACTGG 481

Qy 390 ArgMetTrpAsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCys 408

Db 482 AGAATGTGGAATATGCTGTTTCATAAGTGGATCGTTCGACATATATATTTTCCATGC 538

RESULT 14

AW035727

LOCUS AW035727 561 bp mRNA linear EST 18-MAY-2001

DEFINITION EST281881 tomato callus, TAMU Lycopersicon esculentum cDNA clone

CLC36M23, mRNA sequence.

ACCESSION AW035727

VERSION AW035727.1 GI:5894483

KEYWORDS EST.

SOURCE

ORGANISM

Lycopersicon esculentum

tomato.

REFERENCE

AUTHORS

1 (bases 1 to 561)

Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E.

, Liang, F., Upton, J., Craven, M.B., Bowman, C.L., Ahn, S., Ronning

, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.

Generation of ESTs from tomato callus tissue

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.

Location/Qualifiers

1. 561

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="cLEC36M23"

/clone\_lib="tomato callus, TAMU"

/dev\_stage="25-40 days old"

/tissue\_type="callus"

/lab\_host="XL1-Blue MRF"

/note="vector: pBluescript SK(-); Site1: EcoRI; Site2:

XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons

of seedlings 7-10 days post-germination were excised, cut

at both ends and placed on MS medium with no selection.

Mixed callus was harvested at 25 and 40 days and included

undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 138 a 111 c 118 g 194 t

ORIGIN

Alignment Scores:

Pred. NO.: 1.59e-69 Length: 561

Score: 732.50 Matches: 139

Percent Similarity: 82.00% Conservative: 25

Best Local Similarity: 69.50% Mismatches: 21

Query Match: 27.33% Indels: 15

DB: 10 Gaps: 2

US-09-856-018B-16 (1-504) x AW035727 (1-561)

Qy 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAAlaPheIleValGlu 172

Db 5 CCACCTTCGATGTGCTGTAGTCTCCGATTTCCCTCTTCCCTCTTCTTCTTCGAG 64

Qy 173 LysLeuAlaGlnArgLysCysIleProGluProValValValValHisIleIleIle 192

Db 65 AAAATGGCACAGAAGATATGATGATCAACATGTAGTGTCTTTCACATAATATA 124

Qy 193 ThrSerThrSerLeuPheTyrProValLeuValIleLeuArgCysAspSerAlaPheVal 212

Db 125 ACGACAGTTCATTTTGTATCCAGTTCTGTTCATCCCTCAGGTGTGATCTGCTTTCTA 184

Qy 213 SerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAla 232

Db 195 TCGGTGTACACATGATGTTGCTTGCATGTGTGATGAACTAGTTCTTATGCA 244

Qy 233 HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252

Db 245 CATACAAATTATGATATGACAGAGCTTCCAAAGTCTGTGAATGAGGGTGAG ----- 295

Qy 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272

Db 296 ---AATTCGGAATCAACTACTCTTACAATGTAGTTTTCGAGAGTTTGGCTTACTTCATG 352

Qy 273 ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGly 292

|||||



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 21, 2003, 17:48:52 ; Search time 311 Seconds  
(without alignments)  
3649.541 Million cell updates/sec

Title: US-09-856-018b-16  
Perfect score: 2680  
Sequence: 1 MAISDEPESVATALNHSLSR.....OPMCVLLYYHDLNMRKGLKD 504

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlpl  
-O=/cgn2.1/USPTO\_spool/US09856018/runat\_20022003\_091205\_8317/app\_query.fasta\_1.647  
-DB=N\_Geneseq\_101002 -OFTM=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -SWAP=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09856018 -ECGN\_1.1.396\_erunat\_20022003\_091205\_8317 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2680	100.0	1942	21	AAA48939	Soybean diacylglyc
2	1821	67.9	1904	21	AAA51482	A. thaliana diacyl
3	1821	67.9	1942	21	AAA88835	Arabidopsis acyl C
4	1821	67.9	1942	21	AAZ45371	Acyl-CoA:cholester
5	1821	67.9	1942	22	AAAS0106	Arabidopsis thalia
6	1803	67.3	1888	21	AAA48932	Arabidopsis diacyl
7	1797.5	67.1	1985	21	AAA51484	A. thaliana AS11 d
8	1701.5	63.5	1975	21	AAA48942	Wheat diacylglycer
9	1622	60.5	1587	21	AAA48938	Rice diacylglycer
10	1328	49.6	5193	21	AAA51483	A. thaliana diacyl
11	1321	49.3	1559	21	AAA48935	Corn diacylglycer
12	1319.5	49.2	5339	21	AAA51485	A. thaliana AS11 d
13	1305	48.7	1281	21	AAA48933	Corn diacylglycer
14	1073.5	40.1	901	21	AAA48936	Corn diacylglycer
15	810.5	30.2	978	21	AAA48934	Corn diacylglycer
16	802.5	29.9	1650	21	AAZ49452	Mouse Diacylglycer
17	800.5	29.9	1976	21	AAA76169	Human ACAT related
18	784.5	29.3	1766	21	AAA88846	Rat acyl CoA:chole
19	784.5	29.3	1766	21	AAZ45385	Acyl-CoA:cholester
20	784.5	29.3	1766	22	AAAS01105	Rat sterol acyltra
21	767.5	28.6	1521	19	AAV01533	Human acylcoenzyme
22	739.5	27.6	1895	21	AAA88842	Human acyl CoA:cho
23	739.5	27.6	1895	21	AAZ45383	DNA encoding a pro
24	614.5	22.9	7490	23	ABL28620	Drosophila melanog
25	610	22.8	380	24	ABQ83222	Arabidopsis thalia
26	585	21.8	993	22	ABA08413	Human cDNA SEQ ID
27	568.5	21.2	629	21	AAA51486	EST with homology
28	568.5	21.2	629	21	AAZ49453	A. thaliana Diacyl
29	545.5	20.4	1122	20	AAZ30335	DNA encoding a hum
30	536	20.0	983	19	AAV01539	Human acylcoenzyme
31	492.5	18.4	470	21	AAA48940	Soybean diacylglyc
32	454.5	17.0	3996	23	ABL28621	Drosophila melanog
33	429.5	16.0	452	24	ABL81686	Human ovarian canc
34	421.5	15.7	774	22	AAAS2744	cDNA encoding nove
35	416.5	15.5	470	24	ABL93444	Arabidopsis thalia
36	408.5	15.2	1885	23	ABL13759	Drosophila melanog
37	397	14.8	3649	19	AAV01536	Human acylcoenzyme
38	396.5	14.8	4260	23	ABL13758	Drosophila melanog
39	396	14.8	4011	15	AAQ63212	Acetyl coenzyme A:
40	393.5	14.7	2657	23	ABL15694	Drosophila melanog
41	390	14.6	2040	21	AAA76170	Human ACAT Related
42	389.5	14.5	1509	21	AAZ57360	Human acyl CoA:cho
43	384.5	14.3	542	22	AAAS27552	cDNA encoding nove
44	384	14.3	3650	19	AAAT96368	Human acyl-coenzy
45	376	14.0	1607	21	AAZ57359	Mouse acyl CoA:cho

ALIGNMENTS

RESULT 1  
AAA48939  
ID AAA48939 standard; cDNA; 1942 BP.  
XX  
AC AAA48939;  
XX  
DT 06-DEC-2000 (first entry)  
XX  
DE Soybean diacylglycerol acyltransferase cDNA #1.  
XX  
KW Diacylglycerol acyltransferase; corn; rice; soybean; wheat;  
XX triacylglycerol; herbicide; EC2.3.1.20; ss.  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 29..1543  
/\*tag= a

FT XX /product= Diacylglycerol\_acyltransferase

PN WO200032756-A2.

PD 08-JUN-2000.

PF 01-DEC-1999; 99WO-US28354.

XX 02-DEC-1998; 98US-0110602.

PR 31-MAR-1999; 99US-0127111.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX Cahoon EB, Kinney AJ, Cahoon RE;

XX WPI; 2000-412308/35.

DR P-PSDB; AAY94519.

XX Polynucleotides encoding diacylglycerol acetyltransferase, useful for  
PT synthesis of triacylglycerols and increasing the level of oils in plant  
PT seeds

XX Claim 4; Page 51; 62pp; English.

XX In the present invention, cDNA libraries from Arabidopsis, corn, rice,  
CC soybean and wheat were screened for sequences with homology to a  
CC putative acyl CoA cholesterol acyltransferase related gene from  
CC Arabidopsis thaliana and diacylglycerol acyltransferases from Homo  
CC sapiens and Mus musculus. The cDNA clones identified from this process  
CC were used to form complete diacylglycerol acyltransferase cDNA  
CC sequences. The present sequence is soybean diacylglycerol  
CC acyltransferase cDNA from clone srl.pk0098.a8. Diacylglycerol  
CC acyltransferases are involved in the synthesis of triacylglycerols.  
CC Alteration of the expression of the diacylglycerol acyltransferase  
CC cDNA can be useful for increasing the level of oils in plant seeds.  
CC Inhibitors of diacylglycerol acyltransferase may be useful as  
CC herbicides.

SQ Sequence 1942 BP; 517 A; 425 C; 377 G; 623 T; 0 other;

Alignment Scores:

Pred. No.: 5,17e-275 Length: 1942  
Score: 2680.00 Matches: 504  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 21 Gaps: 0

US-09-856-018b-16 (1-504) x AAA48939 (1-1942)

Qy 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSerLeuArg 20  
Db 29 ATGGCGATTTCCGATGAGCTGAAAGTGTAGCCACTGCTCAACCACTCTTCCTCGGC 88  
Qy 21 ArgArgProSerAlaThrSerThrAlaGlyLeuPheAsnSerProGluThrThrAsp 40  
Db 89 CGCGTCCCTCCGCACTCCACC CGCGGCTCTTCAAATTCGCTGAGACAACCCGAC 148  
Qy 41 SerSerGlyAspAspLeuAlaLysAspSerGlySerAspSerIleAsnSerAspAsp 60  
Db 149 AGTTCGGGTGATGACTGGCCAAAGATTCGGTCCGACGACTCCATCAACAGGAGCAG 208  
Qy 61 AlaAlaValAsnSerGlnGlnAsnGluLysGlnAspThrAspPheSerValLeuLys 80  
Db 209 GCCCGCTCAATTTCCCAACAGCAAAAGGAAACAAAGCAAGCACTGATTTCTCGTCTCAA 268  
Qy 81 PheAlaThrArgProSerValProAlaHisArgLysValLysGluSerProLeuSerSer 100  
Db 269 TTCGCCCTACCGCTCTCCGTCCTCCGCTCAGCGCAAGTGAAGGAAAGTCCGCTCAGTCC 328  
Qy 101 AspThrIlePheArgGlnSerHisAlaGlyLeuPheAsnLeuCysIleValValLeuVal 120  
Db 329 GACACTATTTCCGTCAGAGTCAGCGGGGCTCTTCAACCTTTGTATAGTACTCTTGT 388

Qy 121 AlaValAsnSerArgLeuIleIleGluAsnLeuMetLysTyrGlyTrpLeuIleLysSer 140  
Db 389 GCTGTGAATAGCCGACATCATTTGAGAAATTAATGAGTATGGTGGTGCATCAATCT 448  
Qy 141 GlyPheTrpPheSerSerLysSerLeuArgAspTrpProLeuPheMetCysCysLeuSer 160  
Db 449 GGTCTTTGGTTAGTTCAAAGTCATGAGACATGGCCCTTTTCATGTGTCTTTCT 508  
Qy 161 LeuValValPheProPheAlaPheIleValGluLysLeuAlaGlnArgLysCysIle 180  
Db 509 CTTGTGGTATTTCTCTTCGCTGCTTATAGTGAGAAAGTTGGCACAACGGAAGTGATA 568  
Qy 181 ProGluProValValValLeuHisIleIleThrSerThrSerLeuPheTrpPro 200  
Db 569 CCCGACCAAGTTGTTGTACTTATATATATCATTAATCACTCACTTCGCTTTCTATCCA 628  
Qy 201 ValLeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPhe 220  
Db 629 GTTTTAGTTATTCAGGTGTGATTTCTCTTTTGTATCAGGTGCACGTTAATGCTGTT 688  
Qy 221 SerCysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAla 240  
Db 689 TCTTGTGTGTATGTTAAATTTGGTCTTATGCACATACAACTATGATATGAGACA 748  
Qy 241 LeuThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTrpPro 260  
Db 749 CTTACCAAAATTAGTTGAAAGGGAGAGCACTGCTCGATACTCTGACATGACTATCCT 808  
Qy 261 TyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGln 280  
Db 809 TACAACCTAAGCTTCAAGAGCTTGGCATATTTCTGTTGCTGCTTACATTTATGATCCAG 868  
Qy 281 ProSerTyrProAlaThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeuValLys 300  
Db 869 CCAAGCTATCTCCACACCTTATTTGGAAGGGTGGTGTTCGCCCACTTGTCAAG 928  
Qy 301 LeuIleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIleVal 320  
Db 929 CTGATAATATTTACAGGAGTTATGGGATTTATAATAGACCAATATATTAATCCCATAGTA 988  
Qy 321 GlnAsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLys 340  
Db 989 CAAAATTCACAGCATCTCTCAAGGGAAACCTTCTTACGCCACCCGAGAGAGTCTCGAAG 1048  
Qy 341 LeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrp 360  
Db 1049 CTTTCTGTTCCTCAAAATTTATATGTGTGCTGCGCATGTTCTATTGCTTTTCCACCTTTGG 1108  
Qy 361 LeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrpTrp 380  
Db 1109 TTAATAATCTCTGGCAGAGCTTCTTCGATTTGGTGCATCGTGAATCTACAAGAGTTGGTG 1168  
Qy 381 AsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTrpMet 400  
Db 1169 AATGCCAAAATCTGTGGAAGATTTATGGAGGATGTGGAATATGCTGTGTCAAAATGGATG 1228  
Qy 401 IleArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAlaLeuLeu 420  
Db 1229 ATCGCCACCTATATATTTTCCATGTTTAAGGCACCGCTCTACCAAAAGGCTGCTGCTTTTA 1288  
Qy 421 IleAlaPheLeuValSerAlaLeuPheHisGluLeuCysIleAlaValProCysHisIle 440  
Db 1289 ATTGCCCTCTCTGTTTCTGCTTTTATTCATGAGCTGTCATTTGCTTCTTCCATGACATA 1348  
Qy 441 PheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThrAsn 460  
Db 1349 TTCAAGTTGGGGCTTTTCGTTGGAATATGATGTTTCAGGTTCCTTTGGTCTTGCATCAAT 1408  
Qy 461 TyrLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePhe 480  
Db 1409 TATCTGCAAAATAAATTCAGAAACTCAATGGTGGAAATATGATTTTTTGGTTCATATTC 1468



Qy 353 PheTyrCysPhePheHisLeuTyrLeuAsnTleLeuAlaGluLeuLeuLeuArgPheGlyAsp 372  
Dd 1246 TTCTACTGCTTCTCCACCTTGGTTAAACATATTGGCAGAGCTCTCTGCTCGGGAT 1305  
Qy 373 ArgGluPheTyrLysAspTyrPheAsnAlaLysThrValGluAspTyrTyrArgMetTyr 392  
Dd 1306 CGTGAATTCACAAAGATGTGTGAATGCAAAAGTGTGGAGATTACTGGAGATGTGG 1365  
Qy 393 AsnMetProValHisLysTyrMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412  
Dd 1366 AATATGCTGTTCATAAATGATGGTTCGACATATATATCTCCGTTGCTCGCAGCAAG 1425  
Qy 413 LeuProLysAlaAlaAlaLeuLeuLeuAlaPheLeuValSerAlaLeuPheHisGluLeu 432  
Dd 1426 ATACCAAGACACTCGCCATTATCATCTCTCTAGTCTCTGCAGTCTTTCATGAGCTA 1485  
Qy 433 CysIleAlaValProCysHisIlePheLysLeuTyrPalapheGlyGlyIleMetPheGln 452  
Dd 1486 TGCATCGCAGTTCCTCTCTCTCAAGCTATGGGCTTCTTGGGATTATGTTTCAG 1545  
Qy 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472  
Dd 1546 GTGCTTGTGCTTCATCACAACATACTACAGGAAGTTT---GGCTCAACGTTGGGG 1602  
Qy 473 AsnMetIlePheTyrPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492  
Dd 1603 AACATGATCTCTGTTTCATCTCTGCAATTTTCGACAAACCGATGTGTGCTCTTTAT 1662  
Qy 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503  
Dd 1663 TACCACGACCTGATGAACCGAAAGGATCGATG 1695  
RESULT 3  
AAA88835  
ID AAA88835 standard; DNA; 1942 BP.  
XX  
AC AAA88835;  
XX  
DT 19-FEB-2001 (first entry)  
XX  
DE Arabidopsis acyl CoA:cholesterol acyltransferase DNA.  
XX  
KW Acyl CoA:cholesterol acyltransferase; ACAT; transgenic plant;  
KW sterol; tocopherol; phytosterol; phytostanol; anticholesterolemic;  
KW hypolipemic; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
FH Key Location/Qualifiers  
FT CDS 237..1799  
FT FT /\*tag= a  
FT FT /EC\_number= "2.3.1.26"  
XX  
PN W0200061771-A2.  
XX  
PD 19-OCT-2000.  
XX  
PF 12-APR-2000; 2000WO-US09696.  
XX  
PR 12-APR-1999; 99US-0128995.  
XX  
PA (MONS ) MONSANTO CO.  
XX  
PI Venkatramesh M, Corbin DR, Bhat GB, Boddupalli SS, Grebenok RJ;  
PI Kishore CM, Lardizabal KD, Lassner MW, Rangwala SH, Karunanandaa B;  
XX  
DR WPI; 2000-665136/64.  
DR P-PSDB; AAB19740.  
XX  
PT Genetically engineering the biosynthetic pathways in plants involved in  
PT the accumulation of sterol compounds and tocopherol to produce  
PT compounds for lowering the level of low density lipoprotein cholesterol

in blood serum -

Disclosure; Page 57-58; 166pp; English.

The present sequence is that of Arabidopsis thaliana DNA encoding acyl CoA:cholesterol acyltransferase (ACAT, see AAB19740). Sterol O-acyltransferases such as ACAT catalyze the formation of cholesterol esters from cholesterol and long chain fatty acids. Recombinant constructs of the invention are used to alter the biosynthesis and accumulation of sterols and tocopherols in transgenic plants. Seeds of such plants may contain elevated levels of sitosterol and/or its esters, and alpha-tocopherol, and reduced levels of campesterol and campestanol and their esters. The seeds may also contain the novel sterol brassicasterol. Oil obtained from the seeds can be used in food and pharmaceutical compositions to lower levels of low density lipoprotein cholesterol in blood serum. ACAT enzymes can be used in the present invention to produce elevated levels of phytosterol and/or phytosterol esters.

XX SQ Sequence 1942 BP; 461 A; 421 C; 425 G; 635 T; 0 other;

## Alignment Scores:

Pred. No.:	1.6e-183	Length:	1942
Score:	1821.00	Matches:	353
Percent Similarity:	75.33%	Conservative:	47
Best Local Similarity:	66.48%	Mismatches:	91
Query Match:	67.95%	Indels:	40
DB:	21	Gaps:	9

US-09-856-018B-16 (1-504) x AAA88835 (1-1942)

Qy 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSer----- 18  
Dd 237 ATGGCGATTTTGGATTCT---GCTGGCGTTACTAGCGTGACGGAGACGGTGGCGGAGAG 293  
Qy 19 -----LeuArgArgProSer---AlaThrSerThrAlaGlyLeu 31  
Dd 294 TTCGTCGATCTTGATAGGCTTCGTCGAGGAATCGAGATCGGATTCCTTAACGGACTT 353  
Qy 32 PheAsnSerProGluThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGly 51  
Dd 354 CTTCTCTCTGTTCCGATAATAATCTCTCGGATGATGTTGGAGCTCCCGCCGACCTT 413  
Qy 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67  
Dd 414 AGGATCGGATGATTCGTTGTTAAACGATCAGCTCAGGACAGCAATTTGGCCGGA 473  
Qy 68 GlnAsnGlu-----LysGln 72  
Dd 474 GATAATAACGCTGCTGGCGATAATAACGCTGGTGAAGAGCGCGGAGAGAGAGGA 533  
Qy 73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92  
Dd 534 AACCCGATGCTACG-----TTTACGATCGACCTCGGTTCACGCTCATCGGAGG 584  
Qy 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112  
Dd 585 GCGAGAGAGAGTCCACTTAGCTCCGACCAATCTCAACAGAGCCATGCCGATTATTC 644  
Qy 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMet 132  
Dd 645 AACCTCTGTGTAGTAGTCTTATTCTGCTAAACAGTAGACTCATCATCGAAATCTTATG 704  
Qy 133 LysTyrGlyTyrLeuIleLysSerGlyPheThrPheSerLysSerLysSerLeuArgAspTrp 152  
Dd 705 AAGTATGTTGTTGATCAGCAAGGATTCTGTTTGTAGTTCAGATCGCTCCGAGATTGG 764  
Qy 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172  
Dd 765 CCGCTTTTCATGCTGTGTATATCCCTTTTCGATCTTTCTTTGGCTGCTTTACGGTTGAG 824  
Qy 173 LysLeuAlaGlnArgLysCysIleProGluProValValValValLeuHisIleIleIle 192

```
Db 825 AATGGTACTTCAGAAATACATATACAGAACCTGTTGTCATCTTCTTCATATATTATTC 884
QY 193 ThrSerThrSerLeuPheThrProValLeuValIleLeuArgCysAspSerAlaPheVal 212
Db 885 ACCATGACAGAGGTTTGTATCCAGTTTACCTTAAGCTTAAGGTGTGATCTGCTTTTAA 944
QY 213 SerGlyValThrLeuMetLeuPheSerCysValValThrLeuLysLeuValSerTyrAla 232
Db 945 TCAGGTGTCACTTTGATGCTCTCTCACTTCATGTTGTGTGCTTAAGGTGTGTTCTTATGCT 1004
QY 233 HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252
Db 1005 CATACTAGTCATGACATAGATCCCTAGCCAAATGCAGCTGATAAGGCC----- 1052
QY 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272
Db 1053 -----AATCCTGAAGTCTCTACTACTAGTTAGCTTGAAGAGCTTGGCATATTTTCATG 1103
QY 273 ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGly 292
Db 1104 GTGGCTCCACATGTGTTATCAGCAAGTTATCCACGTTCTGCATGTATACGGAAGGTT 1163
QY 293 TrpLeuPheArgGlnLeuValLysLeuIleLeuPheThrGlyValMetGlyPheIleIle 312
Db 1164 TGGGTGGCTCGTCAATTTCCAAACTGGTCAATATTCACCGGATTCATGGGATTTAATA 1223
QY 313 AspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332
Db 1224 GAACAATATATAAATCTATTGTCAGGAACCTCAAGACATCTTTTGAAGCGCATCTTCA 1283
QY 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352
Db 1284 TATGCTATTGAAGAGTGTGTGAAGCTTTCAGTTCCAAATTTATATGTGTGGCTCTGCATG 1343
QY 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuArgPheGlyAsp 372
Db 1344 TTCTACTGCTTCTCCACCTTGGTTAAACATATTGGCAGAGCTTCTGCTTCGGGGAT 1403
QY 373 ArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392
Db 1404 CGTGAATTCACAAAGATTTGGTGAATGCAAAAGTGTGGGAGATTACCTGGAGATGTGG 1463
QY 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412
Db 1464 ATATGCTGTTTCATAAATGGATGGTTCGACATATATACTTCCCGGCTTGCAGCAAG 1523
QY 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432
Db 1524 ATACCAAGACACTGCCCATATCATGTTCTTCTAGTCTCTGCAGTCTTTCATGAGCTA 1583
QY 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452
Db 1584 TGCATCGCAGTTCCTGTGCTCTCTCAAGCTATGGGCTTTCTTGGGATATATGTTTCAG 1643
QY 453 ValProLeuValLeuIlePheAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472
Db 1644 GTGCTTGGTCTTCATCAACAACATCTACAGAAAGGTTT---GGCTCAACGGTGGGG 1700
QY 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492
Db 1701 ACATGATCTCTCTGTTTCATCTTCGATTTTCGGNACACCGATGTGCTGCTCTTTAT 1760
QY 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503
Db 1761 TACCAGGACCTGATGAACCGAAAGGATCGATG 1793
RESULT 4
AAZ45371
ID AAZ45371 standard; DNA; 1942 BP.
XX
XX AAZ45371;
XX
XX 27-MAR-2000 (first entry)
```

```
XX Acyl-CoA:cholesterol acyltransferase (ACAT)-like protein DNA.
DE
XX Acyl-CoA:cholesterol acyltransferase; ACAT; ACAT-like protein; sterol;
KW ester; triacylglycerol; fatty acyl-CoA; lipid composition; plant cell;
KW diacylglycerol acyltransferase; DAGAT; triglyceride; cancer; diabetes;
KW cardiopulmonary disease; heart failure; atherosclerosis; adipocytosis;
KW leukaemia; skin carcinoma; fibroblastoma; metabolic disorder; obesity;
KW abnormal lipid metabolism; abnormal fat absorption;
KW lipoprotein secretion; adipogenesis; ss.
XX
XX Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 237..1799
FT /*tag= a
FT /product= "Acyl-CoA:cholesterol acyltransferase (ACAT)-
FT like protein"
XX
XX WO963096-A2.
XX
XX 09-DEC-1999.
XX
XX 04-JUN-1999; 99WO-US12541.
XX
XX 05-JUN-1998; 98US-0088143.
XX
XX 12-NOV-1998; 98US-0108389.
XX
XX (CALY ) CALGENE LLC.
XX
XX Lassner MW, Ruezinsky DM;
XX
XX WPI; 2000-105701/09.
XX
XX P-PSDB; AAY54143.
XX
XX Novel polynucleotides used for modifying plant oil composition and for
XX developing products for treating e.g. cancer, diabetes, cardiopulmonary
XX disease or metabolic disorders
XX
XX Claim 4; Fig 1; 89pp; English.
XX
XX The present sequence encodes an acyl-CoA:cholesterol acyltransferase
XX (ACAT) related protein. The ACAT-like protein is active in the formation
XX of a sterol, ester and/or triacylglycerol from a fatty acyl-CoA and
XX sterol and/or diacylglycerol substrate. The DNA can be used for
XX modifying the lipid composition of plant cells. The ACAT-like protein
XX has diacylglycerol acyltransferase (DAGAT) activity, and so the
XX synthesis of triglycerides can be suppressed or increased using the
XX DNA. The protein can be used to produce plant oils with a modified
XX triglyceride content. The products can also be used to identify
XX antagonists and agonists of DAGAT activity. Such agonists and
XX antagonists are particularly useful in treating or ameliorating
XX diseases associated with DAGAT activity, including diseases associated
XX with altered cellular diacylglycerol concentration or PKC activity,
XX atherosclerosis, adipocytosis, leukaemia, skin carcinoma, fibroblastoma,
XX metabolic disorders, obesity, diseases associated with abnormal lipid
XX metabolism, and diseases associated with abnormal fat absorption,
XX lipoprotein secretion and adipogenesis.
XX
XX Sequence 1942 BP; 461 A; 421 C; 425 G; 635 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 1.6e-183 Length: 1942
XX Score: 1821.00 Matches: 353
XX Percent Similarity: 75.33% Conservative: 47
XX Best Local Similarity: 66.48% Mismatches: 91
XX Query Match: 67.95% Indels: 40
XX DB: 21 Gaps: 9
XX
XX US-09-856-018B-16 (1-504) x AAZ45371 (1-1942)
XX
XX 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18
```



Db 237 ATGGCGATTTGGATCTCT---GCTGGCGTTACTACGGTACGGAGAACGGTGGCGGAGAG 293  
 Qy 19 -----LeuArgArgArgProSer---AlaThrSerThrAlaGlyLeu 31  
 Db 294 TTCGTCGATCTTGATAGCTTCGTCGACGGAATTCGAGATCGGATCTTCTAACCGGACTT 353  
 Qy 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGly 51  
 Db 354 CTCTCTCTGCTCGGATAAATAATCTCTTCGGATGATGTTGGAGCTCCCGCGACGCTT 413  
 Qy 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67  
 Db 414 AGGATCGGATGATTCGGTTTAAACGATGACGCTCAGGGAACAGCCAATTTGGCCGGA 473  
 Qy 68 GlnAsnGlu-----LysGln 72  
 Db 474 GATATAACGGTGGTGGCGATAAATAACGGTGGTGAAGAGCGGGGAGAGGAAGAGA 533  
 Qy 73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92  
 Db 534 AACCGGATGCTACG-----TTTACGTATCGACCGTGGTTCACAGCTCATCGGAG 584  
 Qy 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112  
 Db 585 GCGAGAGAGTCCACTTAGCTCCGACGAATCTTCAACACAGAGCCATGCCGGATTATTC 644  
 Qy 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMet 132  
 Db 645 AACCTCTGTGTAGTATCTTATGCTGTAAACAGTAGACTCATCATCGAAATCTTATG 704  
 Qy 133 LysTyrGlyThrLeuIleLysSerGlyPheThrPheSerSerLysSerLeuArgAspTrp 152  
 Db 705 AAGTATGTTGGTGTATCAGAACGATTTCTGGTTTAGTTTCAAGATCGCTCGGAGATTGG 764  
 Qy 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172  
 Db 765 CCGCTTTTCATGTGTGTATATCCCTTCGATCTTTCCTTGGCTGGCTTTACGGTTGAG 824  
 Qy 173 LysLeuAlaGlnArgLysCysIleProGluProValValValValLeuHisIleIleIle 192  
 Db 825 AAATGGTACTTCAGAAATACATATACAGAACTGTGTATCTTCTTCATATATATATC 884  
 Qy 193 ThrSerThrSerLeuPheTyrProValLeuValIleLeuArgCysAspSerAlaPheVal 212  
 Db 885 ACCATGACAGAGGTTTGTATCCAGTTTACGTCACCCCTAAGGTGTGATCTGCTTTTAA 944  
 Qy 213 SerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAla 232  
 Db 945 TCAGTGTCACTTTGATGCTCTCTCACTTGCATTTGTGGCTAAAGTTGGTTTCTTATGCT 1004  
 Qy 233 HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252  
 Db 1005 CATACTAGCTATGACATAAGATCCCTAGCCCAATGCAGCTGATAGGCC-----1052  
 Qy 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272  
 Db 1053 -----AATCTCGAGTCTCTACTACGTTAGCTTTGAAGAGCTTGGCATATTTTCATG 1103  
 Qy 273 ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGly 292  
 Db 1104 CTCGCTCCACATTTGTGTATCAGCCAAATGTTATCCAGTTCTGCATGTATACGGAGGGT 1163  
 Qy 293 TrpLeuPheArgGlnLeuValLysIleIlePheThrGlyValMetGlyPheIleIle 312  
 Db 1164 TGGTGGCTCGTCAATTTGCAAACTGGTCATATTCACCGGATTCATGGGATTTATAATA 1223  
 Qy 313 AspGlnTyrIleAsnProIleValIleAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332  
 Db 1224 GAACAAATATATAATTCCTATTGTGAGGAACCTCAAGCATCTCTTGAAGGCGATCTTCTA 1283  
 Qy 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352

Db 1284 TATGCTATTTGAAGAGAGTGTTCGAAGCTTTCAAGTTCCTCAAAATTTATATGTGTGGCTCTGCATG 1343  
 Qy 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372  
 Db 1344 TTCTACTGCTTCTTCCACCTTTGGTTAAACATATTTGGCAGAGCTTCTCTGCTTCGGGAT 1403  
 Qy 373 ArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392  
 Db 1404 CGTGAATTTCTACAAGATTTGTGGAAATGCAAAAAGTGTGGGAGATTACTGGAGAATGTGG 1463  
 Qy 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412  
 Db 1464 AATATGCTGTTCATAAATGGATGTTCCACATATATATCTCCCGTGTTCGCGAGCAAG 1523  
 Qy 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432  
 Db 1524 ATACAAAGACACTCGCCATTTATCATTTCTTCTTAGTCTCTGCAGTCTTTCATGAGCTA 1583  
 Qy 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452  
 Db 1584 TGCATCGCAGTTCTTGTCTCTTCAAGCTATGGGCTTTTCTTGGGATTTATGTTTCAG 1643  
 Qy 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472  
 Db 1644 GTGGCTTTGGTCTTCATCCAACTATCTACAGGAAGGTTT---GGCTCAACGGTGGGG 1700  
 Qy 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492  
 Db 1701 AACATGATCTTCTGGTTCATCTTCTGCATTTTCGGACAACCGATGTGTGCTCTTTAT 1760  
 Qy 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503  
 Db 1761 TACCAGCACCTGTGTAACCGAAAGGATCGATG 1793  
 RESULT 5  
 AAS01106  
 ID AAS01106 standard; cDNA; 1942 BP.  
 XX  
 AC AAS01106;  
 XX  
 DT 31-MAY-2001 (first entry)  
 XX  
 DE Arabidopsis thaliana sterol acyltransferase ACAT EST sequence.  
 XX  
 KW Lecithin:cholesterol acyltransferase-like; LCAT; sterol acyltransferase;  
 KW acyl CoA:cholesterol acyltransferase-like; ACAT; plant oil; cooking oil;  
 KW nutritional supplement; dairy product; food product; salad dressing;  
 KW expressed sequence tag; EST; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 237..1799  
 FT /\*tag= a  
 FT /product= \*ACAT\*  
 FT  
 PN WO200116308-A2.  
 PD 08-MAR-2001.  
 XX  
 PF 30-AUG-2000; 2000WO-US23863.  
 XX  
 PR 30-AUG-1999; 99US-0152493.  
 XX  
 PA (MONS ) MONSANTO CO.  
 XX  
 PI Lessner M, Van Eenennaam A;  
 XX  
 DR WPI; 2001-169010/17.  
 DR P-PSDB; AAU00462.  
 XX  
 PT New isolated nucleic acid encoding plant lecithin:cholesterol  
 PT acyltransferase-like or acyl (coenzyme A) CoA:cholesterol



PT acyltransferase-like polypeptides, for modifying the sterol content and  
PT oil production of plants -  
PS Claim 4; Page 105-106; 127pp; English.  
XX  
CC The present sequence encodes for Arabidopsis thaliana  
CC acyl CoA:cholesterol acyltransferase-like (ACAT). Several novel  
CC polynucleotides encoding the plant sterol acyltransferases LCAT  
CC (lecithin:cholesterol acyltransferase-like; AAS001081-AAS01104, AAS01341)  
CC and ACAT (AAS01311-AAS01319) are described. A yeast LCAT related open  
CC reading frame, LR01 gene sequence (AAS01342), and a rat ACAT (AAS01105)  
CC cDNA sequence are also described. The polynucleotides encoding LCAT  
CC or ACAT are used to produce LCAT or ACAT polypeptides. They can also  
CC be used in a recombinant construct to transform a host cell (preferably  
CC of a plant) or a plant. The recombinant construct is used to increase or  
CC decrease the sterol content of the host cell or plant. It can be used to  
CC alter oil production of the cell or plant, preferably by increasing it.  
CC The oil of the plant or the plant itself is used as a food product, or  
CC as nutritional or dietary supplements, or in pharmaceutical compositions  
CC for lowering cholesterol. The oil can be used in foods e.g. margarine,  
CC butter, cooking oil, and dressings e.g. salad dressings, mayonnaise,  
CC cheese, processed meat, pasta, sauces, cereals, desserts, dips, chips,  
CC baked goods, pastries, cookies, snack bars, confections, chocolates, and  
CC beverages. The alteration in sterol content and/or composition can also  
CC provide a plant with tolerance to stress and insect damage.  
XX

SQ Sequence 1942 BP; 461 A; 421 C; 425 G; 635 T; 0 other;

Alignment Scores:

Pred. No.: 1.6e-183 Length: 1942  
Score: 1821.00 Matches: 353  
Percent Similarity: 75.33% Conservative: 47  
Best Local Similarity: 66.48% Mismatches: 91  
Query Match: 67.95% Indels: 40  
DB: 22 Gaps: 9

US-09-856-018B-16 (1-504) x AAS01106 (1-1942)

QY 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18  
DB 237 ATGGCGATTGTTGGTCTCT---GCTGGCGTTACTACGGTGACGGAGAACGGTGGCGGAGAG 293  
QY 19 -----LeuArgArgArgProSer---AlaThrSerThrAlaGlyLeu 31  
DB 294 TTCGTCGATCTTGATAGGCTTCTGACCGGAATCAGATCGGATCTCTTAACGGACAT 353  
QY 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspLeuAlaLysAspSerGly 51  
DB 354 CTCTCTCTGGTCCGATAAATCTCTCTCGGATGATGTTGGAGCTCCCGCGACGTT 413  
QY 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67  
DB 414 AGGGATCGGATGATTCCTGTTAACGATGACGCTCAGGGAACAGCCAATTTGGCCGGA 473  
QY 68 GlnAsnGlu-----LysGln 72  
DB 474 GATAATAACGGTGGCGGATTAACGGTGGTGGAGAGCGCGGAGAGAGAGA 533  
QY 73 AspThrAspPheSerValLeuLysPheAlaThrArgProSerValProAlaHisArgLys 92  
DB 534 AACCGCGATGCTACG-----TTTACGTATCGACCGCTCGGTTCCAGCTCATCGGAG 584  
QY 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112  
DB 585 GCGAGAGAGAGTCCACTTACGTCGCGCAATCTTCAACACAGACCATCGCGGATTTATTC 644  
QY 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleGluAsnLeuMet 132  
DB 645 AACCTCTGTGTAGTCTTATGCTGTAAACAGTAGACTCATCATCGAAAATCTTATG 704  
QY 133 LysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrp 152  
DB 705 AAGTATGGTGGTGGATGATCAGACGGATTTCTGGTTAGTTCAAGATCGCTGGCAGATTGG 764

QY 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172  
DB 765 CCGCTTTTCATGTTGTTATATCCCTTTGATCTTCCCTTGGCTGCCCTTACGGTTGAG 824  
QY 173 LysLeuAlaGlnArgLysCysIleProGluProValValValValValLeuHisIleIle 192  
DB 825 AAAATGGTACTTCCAGAAATACATATCAGAACTGTTGTTCATCTTCTTCATATTTATTC 884  
QY 193 ThrSerThrSerLeuPheThrProValLeuValIleLeuArgCysAspSerAlaPheVal 212  
DB 885 ACCATGACAGAGGTTTGTATCCAGTTTACGTCACCCCTAAGCTGTCATCTCTCTTTTA 944  
QY 213 SerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAla 232  
DB 945 TCAGGTGTCTACCTTGTATGCTCTCTACCTTGCATTTGTGTGCTAAAGTTGGTTCTTATGCT 1004  
QY 233 HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252  
DB 1005 CATACTAGCTATGACATAGATCCCTGAGCAATGACGCTGATTAAGGCC----- 1052  
QY 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272  
DB 1053 -----AATCCTGAAGCTCTCCACTAGCTTAGCTTGAAGAGCTTGCATATTTTCATG 1103  
QY 273 ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGly 292  
DB 1104 GTCGCTCCACATGTTGTTATCAGCAAGTTATCCACGTTCTGTCATGTATAGGAAGGT 1163  
QY 293 TrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIle 312  
DB 1164 TGGGTGGCTGCTCAATTTCCAAACTGGTTCATATTCACCGGATTCATGGGATTTATAATA 1223  
QY 313 AspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332  
DB 1224 GAACAATATATAATCTATTGTCAGGAACCTCAAGACATCTTTGAAAGCGCATCTCTTA 1283  
QY 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352  
DB 1284 TATGCTATTGAAAGAGTGTGAAGCTTTCAGTTCCAAATTTATATGTGTGGCTCTGCATG 1343  
QY 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372  
DB 1344 TTCTACTGCTTCTTCCACCTTGGTTAAACATATTTGGCAGAGCTTCTCTGCTTCGGGGAT 1403  
QY 373 ArgGluPheTyrLysAspTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392  
DB 1404 CGTGAATTCACAAAGATTGGTGAATGCAAAAGTGTGGAGATTACTGGAGAATGTGG 1463  
QY 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412  
DB 1464 AATATGCTCTTTCATAAATGGATGGTTCGACATATATACTTCCCGTGTGGCAGCAAG 1523  
QY 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432  
DB 1524 ATACCAAGACACTCCCATTTATCTTCTTCTAGTCTCTGACGCTTCTCATGAGCTA 1583  
QY 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452  
DB 1584 TGCATCGCAGTTCCTTGTCTCTCTCAAGCTATGGCTTTTCTTGGGATATATGTTTCAG 1643  
QY 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472  
DB 1644 GTGCTCTTGTCTTTCATCAAACTATCTACAGGAAGGTTT---GGCTCAACGGTGGGG 1700  
QY 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492  
DB 1701 AACATGATCTTCTGTTTCATCTTCTGCAATTTCCGGAACACCGATGTGTGCTTCTTTAT 1760  
QY 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503  
DB 1761 TACCAGGACCTGATGAACCGAAGGATCGATG 1793

## RESULT 6

AAA48932  
 ID AAA48932 standard; cDNA; 1888 BP.  
 XX  
 AC AAA48932;  
 XX  
 DT 06-DEC-2000 (first entry)  
 XX  
 DE Arabidopsis diacylglycerol acyltransferase cDNA.  
 XX  
 KW Diacylglycerol acyltransferase; corn; rice; soybean; wheat;  
 KW triacylglycerol; herbicide; EC2.3.1.20; ss.  
 XX  
 OS Arabidopsis thaliana.

XX Key Location/Qualifiers  
 FH 131..1693  
 FT CDS  
 FT /\*tag= a  
 FT /product= Diacylglycerol\_acyltransferase

XX WO200032756-A2.

XX 08-JUN-2000.

XX 01-DEC-1999; 99WO-US28354.

XX 02-DEC-1998; 98US-0110602.

XX 31-MAR-1999; 99US-0127111.

XX (DUPO ) DU PONT DE NEMOURS & CO E. I.

XX Cahoon EB, Kinney AJ, Cahoon RE;

XX WPI: 2000-412308/35.

XX P-PSDB; AAY94512.

XX Polynucleotides encoding diacylglycerol acetyltransferase, useful for  
 PT synthesis of triacylglycerols and increasing the level of oils in plant  
 PT seeds -

XX Claim 4; Page 38; 62pp; English.

CC In the present invention, cDNA libraries from Arabidopsis, corn, rice,  
 CC soybean and wheat were screened for sequences with homology to a  
 CC putative acyl CoA cholesterol acyltransferase related gene from  
 CC Arabidopsis thaliana and diacylglycerol acyltransferases from Homo  
 CC sapiens and Mus musculus. The cDNA clones identified from this process  
 CC were used to form complete diacylglycerol acyltransferase cDNA  
 CC sequences. The present sequence is Arabidopsis thaliana diacylglycerol  
 CC acyltransferase cDNA. Diacylglycerol acyltransferases are involved  
 CC in the synthesis of triacylglycerols. Alteration of the expression  
 CC of the diacylglycerol acyltransferase DNA can be useful for  
 CC increasing the level of oils in plant seeds. Inhibitors of  
 CC diacylglycerol acyltransferase may be useful as herbicides.

XX Sequence 1888 BP; 457 A; 394 C; 429 G; 608 T; 0 other;

## Alignment Scores:

Pred. No.: 1-26e-181 Length: 1888  
 Score: 1803.00 Matches: 351  
 Percent Similarity: 74.95% Conservative: 47  
 Best local Similarity: 66.10% Mismatches: 93  
 Query Match: 67.28% Indels: 40  
 DB: 21 Gaps: 9

US-09-856-018b-16 (1-504) x AAA48932 (1-1888)

Qy 1 MetAlaIleSerAspGluProGluSerValAlaThraLeuAsnHisSerSer----- 18  
 |||||  
 Db 131 ATGCGGATTTGGATTCT---GCTGGCGTTACTACGGTGACGGAGAACGGTGGCGGAGAG 187  
 |||||

Qy 19 -----LeuArgArgProSer---AlaThrSerAlaGlyLeu 31  
 |||||  
 |||||

Db 188 TTCTGTCGATCTTGATAGGCTTCGTCGACGGAAATCGAGATCGGATTCCTTAACGGACTT 247  
 Qy 32 PheAsnSerProGluThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGly 51  
 |||||  
 Db 248 CTTCTCTCTGTTCCGATAATAATCTCCTTCGGATGATGTGGAGTCCCGCGAGCTT 307  
 Qy 52 SerAsp-----AspSerIleAsnSerAspAspala-----AlaValAsnSerGlnGln 67  
 |||||  
 Db 308 AGGATCGGATTGATTCGGTTGTTAAAGTACGCTCAGGACACGCAATTTGGCCGGA 367  
 Qy 68 GlnAsnGlu-----LysGln 72  
 |||||  
 Db 368 GATAAATAACGGTGGTGGCGATAATAACGGTGGTGGAGAGCGCGGAGAGAGGA 427  
 Qy 73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92  
 |||||  
 Db 428 AACCCGATGCTAGC-----TTTACGTATCGACCGTGGTTCAGCTCATCGGAGG 478  
 Qy 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112  
 |||||  
 Db 479 GCGAGAGAGAGTCCACTTAGCTCGGACCAATCTTCAACAGAGCATGCGCGATTATTC 538  
 Qy 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleGluAsnLeuMet 132  
 |||||  
 Db 539 AACCTCTGTGTAGTAGTTCTTATTGCTGTAACAGTAGACTCATCATCGAAAATCTTTATG 598  
 Qy 133 LysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrp 152  
 |||||  
 Db 599 AAGTATGTTGGTTGATCAGACGAGATTCTGGTTTAGTTCAGATCGCTCGGAGATTGG 658  
 Qy 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172  
 |||||  
 Db 659 CCGCTTTTCATGTTGATATCCCTTTGATCTTCTTGGTCTTCTTACGGTTGAG 718  
 Qy 173 LysLeuAlaGlnArgLysCysIleProGluProValValValLeuHsIleIle 192  
 |||||  
 Db 719 AAATGTTACTTCAAGAAATACATATCAGACCTGTTGGCATCTTCTTCATATTATTATTC 778  
 Qy 193 ThrSerThrSerLeuPheTrpProValLeuValIleLeuArgCysAspSerAlaPheVal 212  
 |||||  
 Db 779 ACCATGACAGAGGTTTGTATCCAGTTTACGTCACCTCAGGTGATTCGCTTTTAA 838  
 Qy 213 SerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAla 232  
 |||||  
 Db 839 TCAGGTGCTCACTTGTATGCTCCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 898  
 Qy 233 HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252  
 |||||  
 Db 899 CATACTAGTATGACATAGATCCCTAGCCCAATCCAGCTGATAAGGCC----- 946  
 Qy 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272  
 |||||  
 Db 947 -----AATCCTGAAGTCTCCTACTAGCTTAGCTTAGAGACGTTGGCATATTCATG 997  
 Qy 273 ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGly 292  
 |||||  
 Db 998 GTCGCTCCACATTTGTTATTCAGCAAGTATTCAGGCTTCTCCTGATGTATACGGAAGGT 1057  
 Qy 293 TrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIle 312  
 |||||  
 Db 1058 TGGGTGGCTCGTCAATTTGCAAAACTGGTCATATTCACCGGATTCATGGGATTTATAATA 1117  
 Qy 313 AspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyValAsnLeuLeu 332  
 |||||  
 Db 1118 GAACAATATATAATCTCTATGTGTCAGCAAGTATTCAGGCTTCTTGAAGGGGATCTCTA 1177  
 Qy 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352  
 |||||  
 Db 1178 TATGCTATTGAAAGAGTGTGAAGCTTTCAGTTCCAAATTTATATGTGTGCTGCTGATG 1237  
 Qy 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372  
 |||||  
 Db 1238 TTCTACTGCTTCTTCCACCTTTGTTTAAACATATTTGCGAGAGCTTCTCTGCTCGGGAT 1297

QY 373 ArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392  
 DB 1298 CGTGAATTCACAAAGATTGGTGAATGCAAAAGTGTGGAGATTACTGGAGATGTGG 1357  
 QY 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412  
 DB 1358 AATATGCTGTTCATATATGATGGTGTGACATATATATCTCCCGTGTGGCAGCAAG 1417  
 QY 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432  
 DB 1418 ATACCAAGACACTGCCCATATCATGCTTTCCTAGTCTCTGCGAGTCTTTCATGAGCTA 1477  
 QY 433 CysIleAlaValProCysHisIlePheLysLeuTyrAlaPheGlyGlyIleMetPheGln 452  
 DB 1478 TGCATCGCAGTTCCTGTCTCTCAAGCTATGGCTTTCTTGGGATTATGTTTCAG 1537  
 QY 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472  
 DB 1538 GTGCCCTTGTCTTCATCAACAACATCTACAGGAAGGTTT---GGCTCAACGGTGGGG 1594  
 QY 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492  
 DB 1595 AACATGATCTTCTGGTTCATCTCTGCACTTTTCGGACACCGATGTGTCTCTTTAT 1654  
 QY 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503  
 DB 1655 TACCAGACCTGATGAACCGAAAGGATCGATG 1687  
 RESULT 7  
 ID ANA51484 standard; cDNA; 1985 BP.  
 AC ANA51484;  
 DT 09-OCT-2000 (first entry)  
 DE A. thaliana AS11 diacylglycerol acyltransferase cDNA.  
 KW DGAT; diacylglycerol acyltransferase; seed oil; fatty acid synthesis;  
 KW size; weight; carbon flux; TAG1; insertion mutant; ss.  
 OS Arabidopsis thaliana.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 139..1782  
 FT /\*tag= a  
 FT /product= diacylglycerol acyltransferase  
 FT /note= "insertion mutant"  
 FT misc\_RNA 610..690  
 FT /\*tag= b  
 FT /note= "81 bp insertion, duplication of exon 2"  
 XX  
 PN WO200036114-A1.  
 XX  
 PD 22-JUN-2000.  
 XX  
 PF 16-DEC-1999; 99WO-CA01202.  
 XX  
 PR 17-DEC-1998; 98US-0112812.  
 XX  
 PA (CANADA) NAT RES COUNCIL CANADA.  
 XX  
 PI Zou J, Taylor DC, Wei Y, Jako CC;  
 XX  
 DR WPI: 2000-431592/37.  
 DR P-PSDB: AAY96854.  
 XX  
 PT New DNA encoding diacylglycerol acyltransferase from Arabidopsis  
 PT thaliana for transforming plants and regulating seed oil content, fatty  
 PT acid synthesis and seed oil acyl composition in commercial and crop  
 PT plants  
 XX

PS Claim 5; Page 79; 91pp; English.

XX This cDNA encodes Arabidopsis thaliana ecotype Columbia mutant AS11  
 CC diacylglycerol acyltransferase (DGAT). The AS11 mutant TAG1 allele has a  
 CC 147 bp insertion located at the central region of intron 2. The insertion  
 CC is a duplication of a segment that is composed of 12 bp from the 3' end  
 CC of intron 1, the entire sequence of exon 2 (81 bp) and 54 bp from the  
 CC 5' end of intron 2. The DGAT and the insertion mutant (AS11) are useful  
 CC for regulating seed oil content, the ratio of diacylglycerol to  
 CC triacylglycerol proportions in seed oil, fatty acid synthesis, seed oil  
 CC acyl composition, seed size/weight and carbon flux into other seed  
 CC components in commercial and crop plants. The natural formation of  
 CC triacylglycerols can be modified to increase the yield in commercial  
 CC plant oils or modify their composition to achieve specific commercial  
 CC improvements of plants and plant products.

XX Sequence 1985 BP; 482 A; 415 C; 446 G; 642 T; 0 other;

#### Alignment Scores:

Pred. No.: 5.25e-181 Length: 1985  
 Score: 1797.50 Matches: 353  
 Percent Similarity: 71.68% Conservative: 47  
 Best Local Similarity: 63.26% Mismatches: 91  
 Query Match: 67.07% Indels: 67  
 DB: 21 Gaps: 10

US-09-856-018B-16 (1-504) x AAAS1484 (1-1985)

QY 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18  
 DB 139 ATGGCGATTTTGGATTCT---GCTGGCGTTACTACGGTACGGAGAACGGTGGCGGAGAG 195  
 QY 19 -----LeuArgArgArgProSer---AlaThrSerThrAlaGlyLeu 31  
 DB 196 TTCGTCGATCTTGATAGGCTTCGTCGACGAGAAATCGATCGGATCTTCTAACGACTT 255  
 QY 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGly 51  
 DB 256 CTTCTCTCTGTCGCGATAATAATCTCTCTCGGATGATGTTGGAGCTCCCGCCGACGTT 315  
 QY 52 SerAsp-----AspSerIleAsnSerAspAspAla-----AlaValAsnSerGlnGln 67  
 DB 316 AGGGATCGGATTGATCCGTTGTTACGATGACGCTCAGGGAACAGCCCAATTGGCCGGA 375  
 QY 68 GlnAsnGlu-----LysGln 72  
 DB 376 GATAATAACCGTGTGGCGATAATAACGGTGTGGAGAGCGCGGAGAGAGAGCA 435  
 QY 73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92  
 DB 436 AACGCCGATGCTACG-----TTTACGTATCGACCCGCTCGGTTCCAGCTCATCGGAG 486  
 QY 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGln----- 106  
 DB 487 GCGAGAGAGAGTCCACTTAGCTCCGACGCAATCTTCAACACAGCCATCCCGGATTTATTC 546  
 QY 106 ----- 106  
 DB 547 AACCTCTGTGTAGTAGTCTTATTGCTGTAAACAGTAGACTCATCATCGAAATCTTATG 606  
 QY 107 ---SerHisAlaGlyLeuPheAsnLeuCysIleValValLeuValAlaValAsnSerArg 125  
 DB 607 AAGAGCCCATGCCGGATTATTCAACCTCTCTGTAGTAGTCTTATTGCTGTAAACAGTAGA 666  
 QY 126 LeuIleIleGluAsnLeuMetLysTyrGlyTyrLeuIleLysSerGlyPheTrpPheSer 145  
 DB 667 CTCATCATCGAAATCTTATGAGTATGTTGTTGATCGAAGCGGATTTCTGGTTTATG 726  
 QY 146 SerLysSerLeuArgAspTrpProLeuPheMetCysCysLeuSerValValPhePro 165  
 DB 727 TCAAGATCGCTCGGAGATGGCGCTTTTCATGCTGTGTATATATCCCTTCGATCTTCCT 786  
 QY 166 PheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIleProGluProValVal 185



```
QY 143 TrpPheSerSerLysSerLeuArgAspTrpProLeuPheMetCysCysLeuSerLeuVal 162
Db 542 TGGTTTGTAGTCAAGATCGCTGGAGATGGCCACTTCTGATGTGCTCACTTACCC 601
QY 163 ValPheProPheAlaAlaPheIleValGluLeuLysLeuAlaGlnArgLysCysIleProGlu 182
Db 602 ATTTTCCCACTTGTCTCTCATGACGAGAAGTGGGCTCAAGAAGAACTCATCCGTGAT 661
QY 183 ProValValValLeuHisIleIleIleThrSerThrSerLeuPheTyrProValLeu 202
Db 662 CATGTGTCTATTCTTCCATATAATATTACAACCACTGCTTATCTATCCGGTGT 721
QY 203 ValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheSerCys 222
Db 722 GTGATTCTTAAGTGTGAATCAGCAGATATTATCTGGATTGTGTGTTAATGTTCAATGCAAGC 781
QY 223 ValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAlaLeuThr 242
Db 782 ATTACTTGGTGAAGCTTGTCTCTTTGCTCATACAAATATGATATAGGATATTGTCC 841
QY 243 LysLeuValGluLysGlyGlu-----AlaLeuLeuAspThrLeuAsnMetAspTyr 259
Db 842 CAAAGTATTGAAAGGGTCTACACATGCGAGTCTCATCGATGAGGAAACATTAAGGC 901
QY 260 ProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyr 279
Db 902 CCA-----ACTATCAACAGTGTGTGTTATTTTCATGTTGGCCCAACACTTTCTTAC 952
QY 280 GlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeuVal 299
Db 953 CAGCAAGTATCCCGGACAGCATTTATTAGAAAGGCTGGGTACCCGGCAGCTTATA 1012
QY 300 LysLeuIleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIle 319
Db 1013 AAATGTGTAGTTTACAGCTTGATGGGCTTCATAATTCAGCAATACATTAATCAATT 1072
QY 320 ValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeu 339
Db 1073 GTGCAGAAATTCGAAGCATCCATTGAACGGAAATTTCTTGATGCTATTGAGAGAGTCTTG 1132
QY 340 LysLeuSerValProAsnLeuTrpValTrpLeuCysMetPheTyrCysPhePheIleLeu 359
Db 1133 AAATCTCAGTCCCAACATATATGATGGCTTGTATGTTCTATTCCTTTTCCATCAG 1192
QY 360 TrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrp 379
Db 1193 TGGTTCAATATTCTAGCCGAACCTCCCTGCTTTTGGTGTGCTGTAATTCACAAGGACTGG 1252
QY 380 TrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTrp 399
Db 1253 TGGAAAGCCCAACACAGTTGAAGAGTACTGGGAATCTGGAATATGCTGTTCATAAGTGG 1312
QY 400 MetIleArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAlaLeu 419
Db 1313 ATCGTTCGACATATATATTTCCATGCATAGAAGTGGCTTATCAAGGGTGTGCGCAT 1372
QY 420 LeuIleAlaPheLeuValSerAlaLeuPheHisGluLeuCysIleAlaValProCysHis 439
Db 1373 CTCATGCCATTTCTGCTTTTCAGCTGTATTTTCATGAGCTATGATTCGTTCCGTCGCCAC 1432
QY 440 IlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThr 459
Db 1433 ATTTTCAAAATATGGGCATTTTCTGGAATCATGTTTTCAGATTCCCTGCTATTCTTGAGC 1492
QY 460 AsnTyrLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIle 479
Db 1493 AAATATCTTCAGATAGTTCAGAAATACAAATGGTGGGCAACATGATATTTGGTCTTC 1552
QY 480 PheSerIleLeuGlyGlnProMetCysValLeuLeuTyrTyrHisAspLeuMetAsnArg 499
Db 1553 TTCAGCATATGGGCAACCAATGTGTCTTCTTACTACCATCATCATCATGAACAGA 1612
```

```
QY 500 LysGlyLys 502
Db 1613 CAGGCTCAG 1621
RESULT 9
AAA48938
ID AAA48938 standard; cDNA; 1587 BP.
AC AAA48938;
XX 06-DEC-2000 (first entry)
XX Rice diacylglycerol acyltransferase cDNA #2.
XX Diacylglycerol acyltransferase; corn; rice; soybean; wheat;
XX triacylglycerol; herbicide; EC2.3.1.20; ss.
XX Oryza sativa.
XX Key Location/Qualifiers
FH CDS 15..1310
FT /*tag= a
FT /partial
FT /note= "The CDS of this sequence only encodes
XX amino acids 70 to 500 of the protein in AAY94518"
XX WO200032756-A2.
XX 08-JUN-2000.
XX 01-DEC-1999; 99WO-US28354.
XX 02-DEC-1998; 98US-0110602.
XX 31-MAR-1999; 99US-0127111.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX Cahoon EB, Kinney AJ, Cahoon RE;
XX WPI; 2000-412308/35.
XX P-PSDB; AAY94518.
XX Polynucleotides encoding diacylglycerol acetyltransferase, useful for
XX synthesis of triacylglycerols and increasing the level of oils in plant
XX seeds -
XX Claim 4; Page 48-49; 62pp; English.
XX In the present invention, cDNA libraries from Arabidopsis, corn, rice,
XX soybean and wheat were screened for sequences with homology to a
XX putative acyl CoA cholesterol acyltransferase related gene from
XX Arabidopsis thaliana and diacylglycerol acyltransferases from Homo
XX sapiens and Mus musculus. The cDNA clones identified from this process
XX were used to form complete diacylglycerol acyltransferase cDNA
XX sequences. The present sequence is rice diacylglycerol acyltransferase
XX cDNA from clone rls24.pk0034.db.fis. Diacylglycerol acyltransferases are
XX involved in the synthesis of triacylglycerols. Alteration of the
XX expression of the diacylglycerol acyltransferase DNA can be useful for
XX increasing the level of oils in plant seeds. Inhibitors of
XX diacylglycerol acyltransferase may be useful as herbicides.
XX Sequence 1587 BP; 411 A; 323 C; 334 G; 519 T; 0 other;
SQ
```

```
Alignment Scores:
Pred. No.: 1,85e-162 Length: 1587
Score: 1622.00 Matches: 287
Percent Similarity: 82.12% Conservatives: 62
Best Local Similarity: 67.53% Mismatches: 74
Query Match: 60.52% Indels: 2
DB: 21 Gaps: 1
US-09-856-018B-16 (1-504) x AAA48938 (1-1587)
```

```
Qy 75 AspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLysValLys 94
Db 24 GACTTCTCCGG-----TTCCAGCTTCGGCGCGCGCGGTCCAGCGCAAGCGCAAG 77
Qy 95 GluSerProLeuSerSerAspThrPheArgGlnSerHisAlaGlyLeuPheAsnLeu 114
Db 78 GAGAGCCCTTCAGCTCCGAGCCATCTTCAAGCAGAGTCAAGCAGGCTTTTCAACCTA 137
Qy 115 CysIleValValLeuValAlaValAsnSerArgLeuIleLeuGluAsnLeuMetLysTyr 134
Db 138 TGCATGTGTGTTAGTTCAGTCAACAGCAGGCTTATTATCGAGAACTTAATGAAGTAT 197
Qy 135 GlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrpProLeu 154
Db 198 GCCTTATTAAATGAAGAGTGGGTTTTGTTTAAATGATAAATCATTCGCGGACTGGCCACTT 257
Qy 155 PheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGluLysLeu 174
Db 258 CTATGTGTTGCTTAGTCTCCCTGCTTCCCTGGTGGTGCATTTGCAGTTGAAAAGTTG 317
Qy 175 AlaGlnArgLysCysIleProGluProValValValLeuHisIleIleIleThrSer 194
Db 318 GCATTTAAACAATGTTATTACTGATGCTGTTGCTACCTCCCTCCATATCTTCTTCAACA 377
Qy 195 ThrSerLeuPheTyrProValLeuValIleLeuArgCysAspSerAlaPheValSerGly 214
Db 378 ACCGAAATTTGATATCCAGTCTTGCTGATTTCTTAAGTGTGATTCGCAAGTTTGTCTGGC 437
Qy 215 ValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAlaHisThr 234
Db 438 TTTTGTGTGATTTATTTGCTGATGTTGCTGCTGAGCTGTGATCTTTTGGCACATACA 497
Qy 235 AsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeuAspThr 254
Db 498 AACCATGATATAAGCAACTGACCATGGCGGCAAGAGGTTGATATGAAGTAAGCACACA 557
Qy 255 LeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAla 274
Db 558 GTTGACATGGATAATTTACAACTTACAACTTTAGGGAATTAATATATCTCATGATGCT 617
Qy 275 ProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeu 294
Db 618 CCTACATCTGTTATCAGCCAGTATCCCGAACTTCATGCTGTAGAAAAGTTGGCTG 677
Qy 295 PheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIleAspGln 314
Db 678 ATTCTCAAAATATTCTGACTTCTGCTGCTTCTCAAGGCTTCATTATTGAGCAA 737
Qy 315 TyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAla 334
Db 738 TACATAAATCCAATTTGTTGAAATTCAGCATCCATTTGAAAGGAGGACTCCTAAATGCT 797
Qy 335 ThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyr 354
Db 798 GTAGAGACTGTTTGAACACTCTCATTTACCAATGTTTACCTGTGCTTGCATGTTCTAT 857
Qy 355 CysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGlu 374
Db 858 GCTTTTTCATCTCTGCTTAAAGTATCTGCTGAGATCTTCGATTTGGTGACCGTGAA 917
Qy 375 PheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMet 394
Db 918 TTTCACAAAGATTGCTGGAATGCAAAACAAATGATGAGTATTGGAGAAAATGGAATATG 977
Qy 395 ProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGlyLeuPro 414
Db 978 CCTGTACATAAATGGGTGTGCTGCATATTTACTTTCTTGCATGCGAATGGTATATCA 1037
Qy 415 LysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeuCysIle 434
Db 1038 AAGGAAGTTGCTGCTGTGATCATCTCTGTTCTTCCGCTACTCCATGAGATATGTCTC 1097
Qy 435 AlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValPro 454
```

```
Db 1098 GCTGTTCCCTCCGCATTCTCAAGTTCCTGGGCATTCTTAGGAATAATGCTACAGATCCCC 1157
Qy 455 LeuValLeuIleThrAsnTyrLeuGluAsnLysPheArgAsnSerMetValGlyAsnMet 474
Db 1158 CTTATCGTATTGACAGCATACCTCAAAAGTAAATTCAGAGATACAATGGTTGGCAACATG 1217
Qy 475 IlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyrTyrHis 494
Db 1218 ATATTTTGGTTCTTTTCTGCATCTATGGGCAGCAATGTCCCTCTCCTCTACTATCAT 1277
Qy 495 AspLeuMetAsnArg 499
Db 1278 GATGTGATGAACAGG 1292

RESULT 10
AAAS1483
ID AAAS1483 standard; DNA; 5193 BP.
XX
AC AAAS1483;
XX
XX 09-Oct-2000 (first entry)
XX
DE A. thaliana diacylglycerol acyltransferase genomic DNA.
XX
KW DGAT; diacylglycerol acyltransferase; seed oil; fatty acid synthesis;
KW size; weight; carbon flux; ds.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 1761..4780
FT exon /*tag= a 1761..2150
FT exon /*tag= b 1761..2150
FT intron /*partial 2151..2234
FT exon /*tag= c 2235..2315
FT misc_RNA /*tag= d 2235..2315
FT exon /*note= "region duplicated in insertion mutant AS11 (see AAAS1485)" 2235..2315
FT intron /*tag= e 2316..2406
FT exon /*tag= f 2407..2480
FT intron /*tag= g 2481..2582
FT exon /*tag= h 2583..2661
FT intron /*tag= i 2662..2758
FT exon /*tag= j 2759..2823
FT intron /*tag= k 2824..2927
FT exon /*tag= l 2928..3028
FT intron /*tag= m 3029..3108
FT exon /*tag= n 3109..3186
FT intron /*tag= o 3187..3264
FT exon /*tag= p 3265..3366
FT intron /*tag= q 3367..3466
FT exon /*tag= r 3467..3603
FT intron /*tag= s 3604..3747
```

```

FT      exon      /*tag= t      3748..3829
FT      intron    /*tag= u      3830..3933
FT      exon      /*tag= v      3934..3954
FT      intron    /*tag= w      3955..4038
FT      exon      /*tag= x      4039..4101
FT      intron    /*tag= y      4102..4184
FT      exon      /*tag= z      4185..4232
FT      intron    /*tag= aa     4233..4355
FT      exon      /*tag= ab     4356..4418
FT      intron    /*tag= ac     4419..4499
FT      exon      /*tag= ad     4500..4550
FT      intron    /*tag= ae     4551..4675
FT      exon      /*tag= af     4676..4780
FT      exon      /*tag= ag     /partial
XX      XX
PN      WO200036114-A1.
XX      XX
PD      22-JUN-2000.
XX      XX
PE      16-DEC-1999; 99WO-CA01202.
XX      XX
PR      17-DEC-1998; 98US-0112812.
XX      XX
PA      (CANA ) NAT RES COUNCIL CANADA.
XX      XX
PI      Zou J, Taylor DC, Wei Y, Jako CC;
XX      XX
DR      WPI; 2000-431592/37.
DR      P-PSDB; AAY96853.
XX      XX
PT      New DNA encoding diacylglycerol acyltransferase from Arabidopsis
PT      thaliana for transforming plants and regulating seed oil content, fatty
PT      acid synthesis and seed oil acyl composition in commercial and crop
PT      plants
XX      XX
PS      Claim 2; Page 67-68; 91pp; English.
XX      XX
CC      This DNA encodes Arabidopsis thaliana diacylglycerol acyltransferase
CC      (DGAT). The DGAT and an insertion mutant (AS11) are useful for regulating
CC      seed oil content, the ratio of diacylglycerol/triacylglycerol proportions
CC      in seed oil, fatty acid synthesis, seed oil acyl composition, seed
CC      size/weight and carbon flux into other seed components in commercial and
CC      crop plants. The natural formation of triacylglycerols can be modified to
CC      increase the yield in commercial plant oils or modify their composition
CC      to achieve specific commercial improvements of plants and plant products.
XX      XX
SQ      Sequence 5193 BP; 1582 A; 894 C; 903 G; 1814 T; 0 other;

Alignment Scores:
Pred. No.:      2,26e-130      Length:      5193
Score:          1328.00      Matches:      354
Percent Similarity: 39.78%      Conservative: 49
Best Local Similarity: 34.95%      Mismatches: 96
Query Match:      49.53%      Indels:      517
DB:              21      Gaps:      23

US-09-856-018b-16 (1-504) x AAA51483 (1-5193)
QY      1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18

```

```

Db      1761 ATGGCGATTTTGGATTCT---CCTGCGGTTACTACGGTGACGGAACGGTGGCGGAGAG 1817
QY      19 -----LeuArgArgArgProSer---AlaThrSerThrAlaGlyLeu 31
Db      1818 TTTCGTCGATCTTGATAGGCTTCGTCGACGGAATCGAGATCGGATCTTCTCAACGACTT 1877
QY      32 PheAsnSerProGluThrThrThrAspSerSerGlyAspAspLeuAlaLysaspSerGly 51
Db      1878 CTTCTCTCTGGTTCGGATAATAATTCTCTCGGATGATGTTGGAGCTCCGCCGACGTT 1937
QY      52 SerAsp-----AspSerIleAsnSerAspAspAla-----AlaValAsnSerGlnGln 67
Db      1938 ACGGATCGGATGATTCCGTTGTTAAAGATGACGCTCAGGAAACAGCCAATTTGGCCGGA 1997
QY      68 GlnAsnGlu-----LysGln 72
Db      1998 GATAATAACGGTGGTGGCGATAATAACGGTGGTGGAGAGCGCGGAGAGAAGAGA 2057
QY      73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92
Db      2058 AACGCCGATGCTACG-----TTTAGGTATCGACCGTCGGTTCCAGCTCATCGGAGG 2108
QY      93 ValLysGluSerProLeuSerSerAspThrIlePheArg----- 105
Db      2109 GCGAGAGAGAGTCCACTTAGCTCCGACGCAATCTTCAAAACAGGTTTAAATCTCAGAAAT 2168
QY      105 ----- 105
Db      2169 CTTCGAATTTGGTGTGTTGCTGTTTATATGGAATGAGTTGGTGATGTTGTTGCA 2228
QY      106 ----GlnSerHisAlaGlyLeuPheAsnLeuCysIleValValLeuValAlaValAsnSer 124
Db      2229 TTGCAGAGCCATCGCGGATTAATCAACCTCTGTGTAGTAGTTCTTATTGCTGTAACACT 2288
QY      125 ArgLeuIleIleGluAsnLeuMetLys----- 133
Db      2289 AGACTCATCATCGAAATCTTATGAAGGTTTGTGTTACTTGTCTCTCTTTTAGGAATT 2348
QY      134 -----Ty 134
Db      2349 GAATTCCTTGAANAATTTATCAGACGAGTAATACTTTGTTGCTATCATCATGATAGTA 2408
QY      134 rGlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrpProLe 154
Db      2409 TGGTTGGTTGATCAGAAACGGATTCTGTTAGTTCAAGATCGCTGCGAGATTGGCCGCT 2468
QY      154 upheMet----- 156
Db      2469 TTTTCATGTCTGGTAAAGAGATGTTTTTATTTCCAGCAATGTTACATGTTATACGT 2528
QY      157 -----CysCysLeuSe 160
Db      2529 ATAATGATCAGTTTAGTGATCAAGTTCCTTTTGATTCTTCTTCTTGTGGAGTATATC 2588
QY      160 rLeuValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysI 180
Db      2589 CTTTCGATCTTTCTCTTTGGCTGCCTTTACGGTTGAGAAATGGTACTTCAGAAATACAT 2648
QY      180 eProGluProVal-ValValVal----- 187
Db      2649 ATCAGAAACCTGTGATGTAATTAATTTCTCCAGCCATTACTGTAATTTTATTGAAGACAA 2708
QY      188 -----L 188
Db      2709 GTTGTATCATCAGAACTTACAAGTTCGTCTTTGAAATGCTCAAGGTTGTCATCTTTC 2768
QY      188 euHisIleIleIleThrSerThrSerLeuPheTrpProValLeuValIleLeu----- 205
Db      2769 TTCATATTATTATCACCATGACAGAGGTTTGTGTCCAGTTTACGTCACCCCTAAGGTGAT 2828
QY      205 ----- 205

```







CDS 1..984  
 FT /\*tag- a  
 FT /partial  
 FT /product= Partial\_diacylglycerol\_acyltransferase  
 XX PN W0200032756-A2.  
 XX PD 08-JUN-2000.  
 XX PF 01-DEC-1999; 99NO-US28354.  
 XX PR 02-DEC-1998; 98US-0110602.  
 XX PR 31-MAR-1999; 99US-0127111.  
 XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX PI Cahoon EB, Kinney AJ, Cahoon RE;  
 XX DR WPI: 2000-412308/35.  
 XX DR P-PSDB; AAY94515.  
 XX PT Polynucleotides encoding diacylglycerol acetyltransferase, useful for  
 PT synthesis of triacylglycerols and increasing the level of oils in plant  
 PT seeds -  
 XX PS Claim 4; Page 44-45; 62pp; English.  
 XX CC In the present invention, cDNA libraries from Arabidopsis, corn, rice,  
 CC soybean and wheat were screened for sequences with homology to a  
 CC putative acyl CoA cholesterol acyltransferase related gene from  
 CC Arabidopsis thaliana and diacylglycerol acyltransferases from Homo  
 CC sapiens and Mus musculus. The cDNA clones identified from this process  
 CC were used to form complete diacylglycerol acyltransferase cDNA  
 CC sequences. The present sequence is corn diacylglycerol acyltransferase  
 CC cDNA from clone cpjlc.pk005.h23. Diacylglycerol acyltransferases are  
 CC involved in the synthesis of triacylglycerols. Alteration of the  
 CC expression of the diacylglycerol acyltransferase DNA can be useful for  
 CC increasing the level of oils in plant seeds. Inhibitors of  
 CC diacylglycerol acyltransferase may be useful as herbicides.  
 XX SQ Sequence 1559 BP; 413 A; 299 C; 288 G; 555 T; 4 other;  
 Alignment Scores:  
 Pred. No.: 2,06e-130 Length: 1559  
 Score: 1321.00 Matches: 235  
 Percent Similarity: 66.90% Conservative: 54  
 Best Local Similarity: 54.40% Mismatches: 72  
 Query Match: 49.29% Indels: 71  
 DB: 21 Gaps: 1  
 US-09-856-018B-16 (1-504) x AAA48935 (1-1559)  
 QY 142 PheTrpPheSerSerLysSerLeuArgAspTrpProLeuPheMetCysCysLeuSerLeu 161  
 DB 1 TTTTGGTTTAATGCTACATCATGCGAGACTGGCCACTGCTAATGCTGCTTGGCTTACTCTA 60  
 QY 162 ValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIlePro 181  
 DB 61 CCATATTTCCCTTGGTGCATTTGGCAGTCGAAAGTTGGCAATCAACAATCTCGTTAGT 120  
 QY 182 GluProValValValLeuHisIleIleIleThrSerThrSerLeuPheTrpProVal 201  
 DB 121 GATCCGCTGCTACTACCGCTTTTTCACATCCCTTTTACACATTTGAAATTTATATACAGTG 180  
 QY 202 LeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheSer 221  
 DB 181 CTCGTGATTCCTTAAGTGTGATCTGCGACGTTTATACAGGCTTGTGTTGATGTTATTGCC 240  
 QY 222 CysValValTrpLeuLysLeuValSerValAlaHisThrAsnTrpAspMetArgAlaLeu 241  
 DB 241 TGCATTGTTGGCTGAAGCTTGATCTTTTGCACATACAAACCACATCATATAAGAAACATG 300  
 QY 242 ThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTrpProTrp 261

Db 301 ATCACAAAGCGGCAAGAGGTTGATAATGAACTGACCGCGCTGGCATAGATAATTACAA 360  
 QY 262 AsnValSerPheLysSerLeuAlaTrpPheLeuValAlaProThrLeuCysTrpGlnPro 281  
 Db 361 GCTCCAACTCTTGGGAGTCTAACATACTTCAATGATGGCTCCGACACTCTGTTATCAGCCA 420  
 QY 282 SerTrpProArgThrProTrpIleArgLysGlyTrpLeuPheArgGlnLeuValLysLeu 301  
 Db 421 AGTTATCTCGAACACCTTATGTTAGAAAGGTTGGCTGCTCGCTCAAGTATTCTATAC 480  
 QY 302 IleIlePheThrGlyValMetGlyPheIleIleAspGlnTrpIleAsnProIleValGln 321  
 Db 481 TTGATATTTACTGGTCTCCAAAGGATTCATTATTGACCAATACATAAATCTATTGTGTG 540  
 QY 322 AsnSerGlnHisProLeuLysGlyAsnLeuLeuTrpAlaThrGluArgValLeuLysLeu 341  
 Db 541 AACTCTCAACATCCATTGATGGGAGGATTACTGAATGCTGTAGAGACTGTTTTGAAGCTC 600  
 QY 342 SerValProAsnLeuTrpValTrpLeuCysMetPheTrpCysPheHisLeuTrpLeu 361  
 Db 601 TCATTACCAAAATGCTTACCTGGCTTGGCTTTTATTTATTTTCCATCTGTGTTA 660  
 QY 362 AsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTrpLysAspTrpTrpAsn 381  
 Db 661 ACATACTTCTGAGATTTCTCGATTTGGTACCAGGAATTTTACAAAGACTGGTGGAAAT 720  
 QY 382 AlaLysThrValGluAspTrpTrpArgMetTrpAsnMetProValHisLysTrpMetIle 401  
 Db 721 GCAAGACAAATTCATGAGTACTGGACAAATGGAACATGCTGTGCATAAATGGATTGT 780  
 QY 402 ArgHisLeuTrpPheProCysLeuArgHisGlyLeuProLysAlaAlaLeuLeuIle 421  
 Db 781 CGTCATATATATTTCTTTCATCGCAATGTTATATCAAGGAAGTTGCTGTTTATA 840  
 QY 422 AlaPheLeuValSerAlaLeu-Phe- 429  
 Db 841 TGCTTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 429 - 429  
 Db 901 GCATATATTAAATTATATAGTTCTCTATTTTCAAAATGCTGCTTTCGAGTTTCGACATGCT 960  
 QY 429 - 429  
 Db 961 TTTGTTCAAACCTTACCAGCTGTAGATTACTTGGATGAAGTCTCTATATAAAATTCATA 1020  
 QY 430 - 430  
 Db 1021 TTTCACAATCCAGTCCCTTTCGAGAAATATGATACATTTTGTGTCATTTGTACACCA 1080  
 QY 431 uLeuCysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyIleMetPh 451  
 Db 1081 GTTATCGTGTGAGTTCCTCGCCACATCAAGTCTTGGCTTCTTAGGAATCATGCT 1140  
 QY 451 eGlnValProLeuValLeuIleThrAsnTrpLeuGlnAsnLysPheArgAsnSerMetVa 471  
 Db 1141 TCAGATTCCCTCATCATATTGACATCATACCTCAAAAATAAATTCAGTGACACAATGGT 1200  
 QY 471 lGlyAsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLe 491  
 Db 1201 TGCAATATGATCTTTTGGTGTGTTTCTGATATACGGGCGACCAATGTGTGTTCTATT 1260  
 QY 491 uTrpTrpHisAspLeuMetAsnArgLysGlyLys 502  
 Db 1261 GTATTACCATGATGTCATGAACCGGACTGAGAAG 1294  
 RESULT 12  
 ID AAA51485  
 XX AAA51485 standard; DNA; 5339 BP.  
 AC AAA51485;  
 XX

```

09-OCT-2000 (first entry)
A. thaliana AS11 diacylglycerol acyltransferase genomic DNA.
DcAT; diacylglycerol acyltransferase; seed oil; fatty acid synthesis;
size; weight; carbon flux; TAG1; insertion mutant; ds.
Arabidopsis thaliana.

Key Location/Qualifiers
CDS 1761..4926
   FT   /tag= a
   DE   4565..4645
   KW   /tag= b
   KW   2151..2234
   OS   /tag= c
   OS   2235..2315
   OS   /tag= e
   OS   2316..2380
   OS   /tag= f
   OS   2370..2515
   OS   /tag= d
   OS   /note= "insertion region"
   OS   2381..2461
   OS   /tag= g
   OS   /note= "duplication of exon 2"
   OS   2462..2552
   OS   /tag= h
   OS   2553..2622
   OS   /tag= i
   OS   2623..2728
   OS   /tag= j
   OS   2729..2804
   OS   /tag= k
   OS   2805..2901
   OS   /tag= l
   OS   2902..2970
   OS   /tag= m
   OS   2971..3050
   OS   /tag= n
   OS   3051..3174
   OS   /tag= o
   OS   3175..3234
   OS   /tag= p
   OS   3235..3332
   OS   /tag= q
   OS   3333..3410
   OS   /tag= r
   OS   3411..3512
   OS   /tag= s
   OS   3513..3612
   OS   /tag= t
   OS   3613..3749
   OS   /tag= u
   OS   3750..3893
   OS   /tag= v
   OS   3894..3975
   OS   /tag= w
   OS   3976..4079
   OS   /tag= x
   OS   4080..4100
   OS   /tag= y
   OS   4101..4184
   OS   /tag= z
   OS   4185..4247
   OS   /tag= aa
   OS   4248..4330
   OS   /tag= ab
   OS   4331..4378
   OS   /tag= ac
   OS   4379..4501
   OS   /tag= ad

```

```

FT   exon 4502..4564
FT   /tag= ae
FT   intron 4565..4645
FT   /tag= af
FT   exon 4646..4696
FT   /tag= ag
FT   intron 4697..4821
FT   /tag= ah
FT   exon 4822..4926
FT   /tag= ai
FT   /partial
XX   WO200036114-A1.
XX   22-JUN-2000.
XX   16-DEC-1999; 99WO-CA01202.
XX   17-DEC-1998; 98US-0112812.
XX   (CANA ) NAT RES COUNCIL CANADA.
XX   Zou J, Taylor DC, Wei Y, Jako CC;
XX   WPI; 2000-431592/37.
XX   P-PSDB; AAY96854.
XX   New DNA encoding diacylglycerol acyltransferase from Arabidopsis
XX   thaliana for transforming plants and regulating seed oil content, fatty
XX   acid synthesis and seed oil acyl composition in commercial and crop
XX   plants
XX   Disclosure; Page 79-81; 91pp; English.
XX   This DNA encodes Arabidopsis thaliana ecotype Columbia mutant AS11
XX   diacylglycerol acyltransferase (DcAT). The AS11 mutant TAG1 allele has a
XX   147 bp insertion located at the central region of intron 2. The insertion
XX   is a duplication of a segment that is composed of 12 bp from the 3' end
XX   of intron 1, the entire sequence of exon 2 (81 bp) and 54 bp from the
XX   5' end of intron 2. The DcAT and the insertion mutant (AS11) are useful
XX   for regulating seed oil content, the ratio of diacylglycerol to
XX   triacylglycerol proportions in seed oil, fatty acid synthesis, seed oil
XX   acyl composition, seed size/weight and carbon flux into other seed
XX   components in commercial and crop plants. The natural formation of
XX   triacylglycerols can be modified to increase the yield in commercial
XX   plant oils or modify their composition to achieve specific commercial
XX   improvements of plants and plant products.
XX   SQ Sequence 5339 BP; 1620 A; 919 C; 930 G; 1870 T; 0 other;

Alignment Scores:
Pred. No.: 1.9e-129 Length: 5339
Score: 1319.50 Matches: 354
Percent Similarity: 37.98% Conservative: 49
Best Local Similarity: 33.36% Mismatches: 96
Query Match: 49.24% Indels: 565
DB: 21 Gaps: 23

US-09-856-018B-16 (1-504) x AAA51485 (1-5339)
Qy 1 MetaLalSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18
   ||||| |||
Db 1761 ATGGCGATTTTGGATTCT---GCTGGCGTTACTACGGTGACGGAGACGGTGGCGGAGAG 1817
   ||||| |||
Qy 19 -----LeuArgArgProSer---AlaThrSerThrAlaGlyLeu 31
   ||||| ||||| |||
Db 1818 TTCTCGATCTTGATAGCTTTCGTCGAGAAATCGAGATCGGATCTCTTAACGGACTT 1877
   ||||| ||||| |||
Qy 32 PheAsnSerProGluThrThrThrAspSerSerGlyAspLeuAlaLysAspSergly 51
   ||||| ||||| |||
Db 1878 CTTCTCTCGTCCGATATAATTCCTTCGGATGATGTTGGNGCTCCCGCGACGTT 1937
   ||||| ||||| |||
Qy 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67

```



QY	395	-----ProValHisIysTrpMetIleA	402
Db	4147	AAGAAGACAGCTAATTAAGATCTGTTTCCCTTGACACGCCCTGTTTCATAAAATGGATGTTTC	4206
QY	402	rgHisLeuTyrrPheProCysLeuAtrgHisIleGlyLeuProLysAlaAla	417
Db	4207	GACATATATACTTCCCGTCTGCTGCGCAGCAGATACCAAGGTGAG-TGAGATATATACC	4265
QY	417	-----	417
Db	4266	GATATGCAATTGTCGAGATTTGTTTCTGTGATATATAAATTTAACCCCTCCACACACTTGT	4325
QY	418	-----AlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGlu	431
Db	4326	TTCAGACACTCGCCATATCATCTTCTTCTAGTCTCTGCGAGTCTTTCATGAGGTATACA	4385
QY	431	-----	431
Db	4386	TACTTCTACATTGCCCTGTCTCTAGAGCATGAACACACACGCTAGTGAAGAANAATGCTAA	4445
QY	432	-----LeuC	433
Db	4446	TATTAAGCATTTGTTTCTTACTTAACGATCTTGCTGTACAAATTCCTTTTGACAGCAT	4505
QY	433	ysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyIleMetPheGln	452
Db	4506	GCATCGCAGTTCCTTCTGCTCTCTCAAGCTATGGCTTTCTTGGGATATGTTTCAGG	4565
QY	452	-----	452
Db	4566	TTAAAAAATTACTAAACTGCTGCAGTCGATTTTACTAACTCTAATCTCATATTTCTGAC	4625
QY	453	-----ValProLeuValIleThrAsnTyrrIleuGlnAsnLysp	466
Db	4626	CAACCAATTTGTTGAGTAGGTGGCTTTGGTCTTCATCACAAATCTCTACAGGAAGGT	4685
QY	466	he-ArgAsnSerMet	470
Db	4686	TTGGCTCAACGGTATGCTCTCAAAACCCGAGAAAATAGAACGAATAACTCTTCTTTCAT	4745
QY	470	-----	470
Db	4746	AGCTAGCCATTTAAATCCGAATGCTGAACCTTAATAATAAAGGTGATCTGTTTGGAA	4805
QY	471	-----ValGlyAsnMetIlePheTrpPheIlePheSerIleLeuGlyG	485
Db	4806	GGGATCATATTATTAGTGGGGAACATGATCTCTGGTTCATCTTCGATTTTCGACA	4865
QY	485	nProMetCysValLeuLeuTyrrTyrrHisAspLeuMetAsnArgLysGlyLysLeu	503
Db	4866	ACCGATGTGTGCTTCTTTTATTACACGACCTGATGAACCGAAAGGATCGCATG	4920
RESULT	13		
AAA48933			
ID	AAA48933	standard; cDNA; 1281 BP.	
XX	AAA48933;		
XX	06-DEC-2000	(first entry)	
DT			
XX		Corn diacylglycerol acyltransferase cDNA #1.	
DE			
XX		Diacylglycerol acyltransferase; corn; rice; soybean; wheat;	
KW		triacylglycerol; herbicide; EC2.3.1.20; ss.	
XX			
OS		zea mays.	
XX			
PH		Location/Qualifiers	
FT		1..1090	
FT		/tag= a	
FT		/partial	
FT		/transl_except= (pos:283..286,aa:ile)	

[illegible]

Db 241 GTTGGCTGACGTTGTATCTTTTGCACATACAAACCATGATATATAAGGAAACAGTATAC 300  
 QY 244 LeuValGluLysGly-GluAlaLeuLeuAspThrLeuAsnMetAspTyrProTyrAsnVa 263  
 Db 301 AAGCGCAAGAGGTTGATAATGAATGACCGCGGTGCATAGATAAATTTACAANCTCC 360  
 QY 263 lSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGln-Pro-Ser 282  
 Db 361 AACTCTTGGGATCTCAACATCTCAAGATGGCTCCGACACTCTGTATCAAGCCAAAGT 420  
 QY 283 TyrPro-ArgThrProTyrIleArgLysGlyTyrPheLeuAspThrLeuValLysLeu1 302  
 Db 421 TATCTNCGAACACCTTATGTAGAAAGTTGGTGGTCCGTCAGATTATTTCTACTT 480  
 QY 302 ellePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIleValGlnAs 322  
 Db 481 GATATTTACTGCTCTCCAAAGGATTCATTTATTGAGCAATACATAAATCCCTATTGTGAA 540  
 QY 322 nSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSe 342  
 Db 541 CTCTCAACATCCATTGATGGGAGGATTTACTGAATGCTGTAGAGACTGTTTTGAAGCTTC 600  
 QY 342 rValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeuAs 362  
 Db 601 ATTACCAATGCTACCTGCTGGCTTGCATGTTTATTGCTTCCATCTCTGGTTAA 660  
 QY 362 nIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTyrTrpAsnAl 382  
 Db 661 CATACTTGTGATCTCTTCGATTTGGTGACCGAGAAATTTCTACAAGACTGGTGGAAATGC 720  
 QY 382 aLysThrValGluAspTyrTrpArgMetTrpAsnMetProValHisLysTrpMetIleAr 402  
 Db 721 AAGACATTTGATGAGTACTGAGAAATGGAACATGCTGTCATTAATGATGTTGTCG 780  
 QY 402 gHisLeuTyrPheProCysLeuArgHisGlyLeuProLysAlaAlaLeuLeuIleAl 422  
 Db 781 TCATATATATTTCCCTTGCATCGGAATGATATCAAGGAAGTTGCTGTTTATATATC 840  
 QY 422 aPheLeuValSerAlaLeuPheHisGluLeuCysIleAlaValProCysHisIlePheLy 442  
 Db 841 GTTCTTTGTTCTGCTGTTACTTATGAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 900  
 QY 442 sLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeuValLeuIleThrAsnTyrLe 462  
 Db 901 GTTCTGGGCTTCTTAGGAATCATGCTCAGATTCCTTCATCATATATGACATCATCT 960  
 QY 462 uGlnAsnLysPheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSerI1 482  
 Db 961 CAAAATAAATTCAGTGACACAATGGTTGGCAATATGATCTTTTGGTTTTTCTGCAT 1020  
 QY 482 eLeuGlyGlnProMetCysValLeuLeuTyrTyrHisAspLeuMetAsnArgLysGlyLy 502  
 Db 1021 ATACGGGCGCAATGTGTCTATTGTATTATGATGATGATGATGATGATGATGATGATGATG 1080  
 QY 502 s 502  
 Db 1081 G 1081  
 RESULT 14  
 AAA48936  
 ID AAA48936 standard; cDNA; 901 BP.  
 XX  
 AC AAA48936;  
 DX  
 TT 06-DEC-2000 (first entry)  
 DE  
 DE Corn diacylglycerol acyltransferase cDNA #4.  
 XX  
 KW Diacylglycerol acyltransferase; corn; rice; soybean; wheat;  
 KW triacylglycerol; herbicide; EC2.3.1.20; ss.  
 XX  
 OS Zea mays.  
 XX

Key Location/Qualifiers  
 CDS 1..858  
 /tag= a  
 /partial  
 /transl\_except= (pos:442..444,aa:Xaa)  
 /transl\_except= (pos:490..492,aa:Xaa)  
 /transl\_except= (pos:587..589,aa:Xaa)  
 /transl\_except= (pos:805..807,aa:Xaa)  
 /transl\_except= (pos:820..822,aa:Xaa)  
 /note= "Xaa= unknown"  
 /product= Partial\_diacylglycerol\_acyltransferase  
 WO200032756-A2.  
 XX  
 PN 08-JUN-2000.  
 XX  
 PD 01-DEC-1999; 99WO-US28354.  
 PF  
 XX 02-DEC-1998; 98US-0110602.  
 PR 31-MAR-1999; 99US-0127111.  
 XX  
 PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
 XX  
 PI Cahoon EB, Kinney AJ, Cahoon RE;  
 XX  
 DR WPI; 2000-412308/35.  
 DR P-PSDB; AAY94516.  
 XX  
 PT Polynucleotides encoding diacylglycerol acetyltransferase, useful for  
 PT synthesis of triacylglycerols and increasing the level of oils in plant  
 XX seeds -  
 PS Claim 18; Page 46; 62pp; English.  
 XX  
 CC In the present invention, cDNA libraries from Arabidopsis, corn, rice,  
 CC soybean and wheat were screened for sequences with homology to a  
 CC putative acyl CoA cholesterol acyltransferase related gene from  
 CC Arabidopsis thaliana and diacylglycerol acyltransferases from Homo  
 CC sapiens and Mus musculus. The cDNA clones identified from this process  
 CC were used to form complete diacylglycerol acyltransferase cDNA  
 CC sequences. The present sequence is corn diacylglycerol acyltransferase  
 CC cDNA from a contig of clones p0042, cspaf49, p0122 ckamb57r and  
 CC p0125.czaau61b. Diacylglycerol acyltransferases are involved in the  
 CC synthesis of triacylglycerols. Alteration of the expression of the  
 CC diacylglycerol acyltransferase DNA can be useful for increasing the  
 CC level of oils in plant seeds. Inhibitors of diacylglycerol  
 CC acyltransferase may be useful as herbicides.  
 XX  
 SQ Sequence 901 BP; 256 A; 172 C; 193 G; 276 T; 4 other;  
 Alignment Scores:  
 Pred. No.: 2,07e-104 Length: 901  
 Score: 1073.50 Matches: 194  
 Percent Similarity: 82.73% Conservative: 36  
 Best Local Similarity: 69.78% Mismatches: 40  
 Query Match: 40.08% Indels: 8  
 DB: 21 Gaps: 3  
 US-09-856-018B-16 (1-504) x AAA48936 (1-901)  
 QY 229 ValSerTyrAlaHisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGly 248  
 Db 31 GTCTTTATGCACATACAAATATGATATATGAGGTATGTCACAAAGTACTGAGAAGGT 90  
 QY 249 GluAla-----LeuLeuAspThrLeuAsnMetAspTyrProTyrAsnValSerPhe 265  
 Db 91 GCTGCATATGGAATATATCTGATCTCGATATGAAATGCA-----ACCTTT 141  
 QY 266 LysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGlnProSerTyrProArg 285  
 Db 142 AAAAGTCTAGTGTACTTCTATGTTGGCCCCAACACTTTGTTACCGACCACTTATCTCA 201  
 QY 286 ThrProTyrIleArgLysGlyTrpLeuPheArgGlnLeuValLysLeuIlePheThr 305

```
Db 202 ACTACATGTAATAGAAAGGGTGGTGACCGACCAACTATAAAGTGGCTGGTTTACAA 261
Qy 306 GlyValMetGlyPheIleAspGlnTyrIleAsnProIleValGlnAsnSerGlnHis 325
Db 262 GGCCTGATGGCTTCATTAATGAGCAATATATAAACCACCAATGTGAAGAATCCCAACAT 321
Qy 326 ProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSerValProAsn 345
Db 322 CCACCTGAAGGAATTTTGAATGCTATAGAAAGAGCTTTAAAACTCTCAGTCCCAACA 381
Qy 346 LeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTyrLeuAsnIleLeuAla 365
Db 382 TTATATGATGGCTTTCATGCTTCTATGCTTTTTCATTTATGCTGGAACATGTAGCT 441
Qy 366 GluLeuLeuArgPheGlyAspArgGluPheTyrIleAspTrpTrpAsnAlaLysThrVal 385
Db 442 TAACCTCCTCTGTTTCGGTGACCGGAAATCTATAAGGACTGGTGAATGNCAAACATGTT 501
Qy 386 GluAspTyrTrpArgMetTrpAsnMetProValHisLysTrpMetIleArgHisLeuTyr 405
Db 502 GAAGAGTACTGGAGAGTGTGGAACATGCTGTTTCATTAAGTGGATCATCAGACATATAT 561
Qy 406 PheProCysLeuArgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuVal 425
Db 562 TTTCATGATTAAGGNAAGGCTTTTCCAGGGGTGATGCTATCTTAATCTCGTTTCTGTT 621
Qy 426 SerAlaLeuPheHisGluLeuCysIleAlaValProCysHisIlePheLysLeuTrpAla 445
Db 622 TCAGCTGATTCATGATAGATATGATTCGGTGGCTGCCACATTTCAAAATTCGGCA 681
Qy 446 PheGlyGlyIleMetPheGlnValProLeuValLeuIleThrAsnTyrLeuGlnAsnLys 465
Db 682 TTTTCTGGGATCATGTTTCAGATACCGTGTGATTTCTTGACAAGATATCTCCATGCTACG 741
Qy 466 PheArgAsnSerMetValGlyAsnMetIlePheTrpPheIlePheSerIleLeuGlyCln 485
Db 742 TTCAGCATGTAAATGGGGGCAATGATATTTGGTTC---TTCAGTATAGTCCGACAG 798

RESULT 15
AAA48934
ID AAA48934 standard; cDNA; 978 BP.
XX
AC AAA48934;
XX
DT 06-DEC-2000 (first entry)
XX
DE Corn diacylglycerol acyltransferase cDNA #2.
XX
KW Diacylglycerol acyltransferase; corn; rice; soybean; wheat;
KW triacylglycerol; herbicide; EC2.3.1.20; ss.
XX
OS Zea mays.
XX
FH Key 3..470
FT CDS /*tag= a
FT /*codon_start= 3
FT /partial
FT /product= Partial_diacylglycerol_acyltransferase
XX
WO200032756-A2.
XX
PD 08-JUN-2000.
XX
PF 01-DEC-1999; 99WO-US28354.
XX
PR 02-DEC-1998; 98US-0110602.
PR 31-MAR-1999; 99US-0127111.
```

```
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX Cahoon EB, Kinney AJ, Cahoon RE;
XX
DR WPI: 2000-412308/35.
DR P-PSDB: AAY94514.
XX
PT Polynucleotides encoding diacylglycerol acetyltransferase, useful for
PT synthesis of triacylglycerols and increasing the level of oils in plant
PT seeds
XX
PS Claim 15; Page 43; 62pp; English.
XX
CC In the present invention, cDNA libraries from Arabidopsis, corn, rice,
CC soybean and wheat were screened for sequences with homology to a
CC putative acyl CoA cholesterol acyltransferase related gene from
CC Arabidopsis thaliana and diacylglycerol acyltransferases from Homo
CC sapiens and Mus musculus. The cDNA clones identified from this process
CC were used to form complete diacylglycerol acyltransferase cDNA
CC sequences. The present sequence is corn diacylglycerol acyltransferase
CC cDNA from clone cen3n.pk0113.e12. Diacylglycerol acyltransferases are
CC involved in the synthesis of triacylglycerols. Alteration of the
CC expression of the diacylglycerol acyltransferase DNA can be useful for
CC increasing the level of oils in plant seeds. Inhibitors of
CC diacylglycerol acyltransferase may be useful as herbicides.
XX
SQ Sequence 978 BP; 287 A; 173 C; 182 G; 336 T; 0 other;

Alignment Scores:
Pred. No.: 2.4e-76 Length: 978
Score: 810.50 Matches: 142
Percent Similarity: 70.28% Conservative: 33
Best Local Similarity: 57.03% Mismatches: 24
Query Match: 30.24% Indels: 50
DB: 21 Gaps: 2

US-09-856-018B-16 (1-504) x AAA48934 (1-978)
Qy 289 IleArgLysGlyTrpLeuPheArgGlnLeuValLysLeuIlePheThrGlyValMet 308
Db 9 GTTAGAAGAAGTGGCTGGTCCGTCAGGTATTCTATCTATGATATTACTGGTCTCCAA 68
Qy 309 GlyPheIleIleAspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLys 328
Db 69 GGATTCATATTAGCAATACATAAATCCTATTGTTGGAACCTCAACATCCATTTGATG 128
Qy 329 GlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrVal 348
Db 129 GGAGGATTACTGAATGCTGTAGAGACTGTTTGAAGCTCTCATCCCAAAATGTCTACCTG 188
Qy 349 TrpLeuCysMetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeu 368
Db 189 TGGCTTTGCATGTTTATTCGCTTTTCATCTGTTGTTAAACATACACTGCTGAGATTCCT 248
Qy 369 ArgPheGlyAspArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyr 388
Db 249 CGATTTGGTGACCGAGAAATTCACAAAGACTGTGTGAATGCAAAAGACAATTTGATGACTAC 308
Qy 389 TrpArgMetTrpAsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCys 408
Db 309 TGGAGAAATGGAACATGCCTGTGCATAAATGGAATGTTTCGTCATATATATTTCCCTTGC 368
Qy 409 LeuArgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeu 428
Db 369 ATGCGAAATGGTATATCAAGGAAGTTCGCTGTTTATATGCTTCTTCTTCTTCGCTGTA 428
Qy 429 -Phe-
Db 429 CTTTCATGAGTGCAGATTACTTTGATGAAGTGCTCTATATAAAATTAATATTTTATAAT 488
Qy 430 -----HisGluLeuCysIle 434
|||||
```

```
Db 489 CCAGTCCCTTCGAGAAATTATGATACATTTTGTTCGAATTTGTACACCAAGTTATGTGT 548
Qy 434 ealaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValPr 454
Db 549 TGCCTGTTCCCTGCCACATACTCAAGTCTCGGCTTTCTTAGGAATCATGCTTCAGATTCC 608
Qy 454 oLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMe 474
Db 609 CCTCATATATTGACATCATACCTCAAAATAAATTCAGTGACACAATG----- 657
Qy 474 tilePheTrpPheIlePheSerIleLeuGlyGlnPrometCysValLeuLeuTyrTyrHI 494
Db 658 -----CCAAATGTGTGTTCTTATTGTATTACCA 683
Qy 494 sAspLeuMetAsnArgLysGlyLys 502
Db 684 TGATGTGATGAACCGGACTGAGAAG 708
```

Search completed: February 21, 2003, 19:51:58  
Job time : 348 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 21, 2003, 18:31:35 ; Search time 3369 Seconds  
(without alignments)  
4353.759 Million cell updates/sec

Title: US-09-856-018b-16

Perfect score: 2680

Sequence: 1 MAISDEPESVATALNHSSLR.....QPMCVLLYYHDLNMRKGKLD 504

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEVEL=xl  
-O=/cn2\_1/USPTO\_spool/US09856018/runat\_20022003\_091205\_8339/app\_query.fasta\_1.647  
-DB=GenEmbl -OPWT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09856018@cgn\_1\_1\_3745-brunat\_20022003\_091205\_8339 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAR -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

29: em.vi:\*  
30: em.htg\_hum:\*  
31: em.htg\_inv:\*  
32: em.htg\_other:\*  
33: em.htg\_mus:\*  
34: em.htg\_pln:\*  
35: em.htg\_rod:\*  
36: em.htg\_mam:\*  
37: em.htg\_vrt:\*  
38: em.sy:\*  
39: em.htg\_hum:\*  
40: em.htg\_mus:\*  
41: em.htg\_others:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1943	72.5	1964	8	AF298815 Perilla f
2	1871	69.8	2099	8	AF129003 Nicotiana
3	1861	69.4	1537	8	AF251794 Brassica
4	1836	68.5	2090	8	AY084052 Tropaeolu
5	1821	67.9	1845	8	AF051849 Arabidops
6	1821	67.9	1904	8	ATH238008 Arabidops
7	1821	67.9	1942	6	AX090345 Sequence
8	1821	67.9	1942	6	AX090349 Sequence
9	1821	67.9	1988	8	ATH131831 Arabidops
10	1821	67.9	2005	8	AY054480 Arabidops
11	1794.5	67.0	1512	8	AF164434 Brassica
12	1519	56.7	1446	8	AF155224 Brassica
13	1328	49.6	92822	8	AC005317 Arabidops
14	1328	49.6	97495	8	AC003058 Arabidops
15	807.5	30.1	1650	10	BC003717 Mus muscu
16	807.5	30.1	1776	10	BC003717 Mus muscu
17	800.5	29.9	1976	9	AF059202 Homo sapi
18	800.5	29.9	1998	9	BC015762 Homo sapi
19	795.5	29.7	1792	9	AF236018 Cercopith
20	784.5	29.3	1766	6	AX090340 Sequence
21	779.5	29.1	1751	10	AF296131 Rattus no
22	769.5	28.7	1935	4	AY093657 Sus scrof
23	764.5	28.5	1493	9	BC006263 Homo sapi
24	753	28.1	185932	2	AP003714 Oryza sat
25	752	28.1	1497	3	AF221132 Caenorhab
26	739.5	27.6	1895	6	AX090339 Sequence
27	739.5	27.6	2074	3	AY051835 Drosophil
28	734.5	27.4	2117	3	AF468650 Drosophil
29	734.5	27.4	2342	3	AF468649 Drosophil
30	678.5	25.3	34948	3	CBRG33E23
31	632	23.6	11966	3	CEH19N07
32	614.5	22.9	40494	2	AC005122
33	614.5	22.9	60144	2	AC019798
34	614.5	22.9	91685	2	AC005121
35	614.5	22.9	170362	3	AC009250
36	614.5	22.9	258166	3	AE003652
37	570	21.3	728	10	RNO345014
38	546.5	20.4	253305	3	PFMAL3P7
39	420	15.7	262797	3	CNS07EGG
40	410	15.3	8593	4	AY065621 Bos tauru
41	408.5	15.2	2102	3	AY113537 Drosophil
42	406	15.1	1750	10	D86373 Rattus norv
43	405	15.1	645	11	PMIC28
44	400	14.9	3041	10	S81092
45	400	14.9	3697	10	MUSACACT

#### ALIGNMENTS

RESULT 1

```

AF298815      1964 bp      mRNA      linear      PLN 16-OCT-2000
LOCUS      Perilla frutescens diacylglycerol acyltransferase (Dgat1) mRNA,
DEFINITION      complete cds.
ACCESSION      AF298815
VERSION      AF298815.1 GI:10803052
KEYWORDS
SOURCE      Perilla frutescens.
ORGANISM
Perilla frutescens.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.
REFERENCE
AUTHORS      Huang,S.-K. and Hwang,Y.-S.
TITLE      Isolation of Perilla frutescens diacylglycerol acyltransferase cDNA
JOURNAL
REFERENCE      2 (bases 1 to 1964)
AUTHORS      Huang,S.-K. and Hwang,Y.-S.
TITLE      Direct Submission
JOURNAL
FEATURES
source      1..1964
/organism="Perilla frutescens"
/cultivar="Oxdong"
/db_xref="taxon:48386"
1..1964
/gene="Dgat1"
69..1673
/gene="Dgat1"
/codon_start=1
/note="Dgat"
/product="diacylglycerol acyltransferase"
/protein_id="AAG23696.1"
/db_xref="GI:10803053"
/translaton="MAILDSPEILDTSSADNGAAHTTLRRQSRVPPPLDSDS
NSLEASINDENVRNANLIENRGAVSENEKQESYGEKGAKEKNGETSGN
GTDVMAKFTFRAAPAHKRNKESPLSSDAIFKOSHAGLFNLCIVLVAVNSRLIEN
LMRYGMLKSGFWSFTSLRDWPLLMCCLSLFPALASFLVEKLVLKLYIPWVAVFL
HVTITTEILFPVVDYILRCDASVLSVTLMLFACTVWLKLVSAHTNVDRLVLAESLD
KWMASRYNLDYVDYDFKSLAYFVAVPTLCYQPSVPRTACIRKGVVRLKLVIF
TLGLGFTIOYIPINQVONSOHLKGLNLYAIEVLKLSVPLNLYWLCMFCFPHLNLN
ILALLCFGDREYKQDWNARTVEEVRWMMNPVHKWVRHITCPCLQNGIPKIVAVL
IALVSAIFHELCLVAVPCQIFKFWAFSGIMLQVPLVIVTILQEKFNKSMVGNMFWC
FCICFGCPMCLVILYHDLNMRKASAR"
BASE COUNT      522 a 395 c 461 g 586 t
ORIGIN
Alignment Scores:
Pred. No.:      1..51e-162      Length:      1964
Score:      1943.00      Matches:      369
Percent Similarity:      77.74%      Conservative:      50
Best Local Similarity:      68.46%      Mismatches:      72
Query Match:      72.50%      Indels:      48
DB:      8      Gaps:      6

US-09-856-018b-16 (1-504) x AF298815 (1-1964)
Qy      1 MetAlaIleSerAspGluProGluSerValAlaThr-----12
Db      69 ATGGCGATCTTGGACTCGCGGAGATCCTGGACACAGCGTCCGTCAGTGGCGGACACGGC 128
Qy      13 AlaLeuAsnHisSerSerLeuArgArgProSerAlaThrSerThrAlaGlyLeuPhe 32
Db      129 GCCGCATCACACCACTCTTCGCGGAGACAAAGTCGCGCTCCGTTCCGCCCTCTTCTC 188
Qy      33 AsnSerProGluThrThrAspSerSerGlyAspPheLeuAlaLysAspSerGlySer 52
Db      189 -----GATCCGATTCACACTCTCTGTGAGGAGAGAGCGCAATC 227
Qy      53 AspAspSerIleAsn-----SerAspAspAla-----Ala 62
Db      228 AATGATCCGGAATATGTCGAACAGACGCTAATTTGATCGAANAATCTCCGCGGCGGAGCC 287

```

```

Qy      63 ValAsnSerGlnGlnGlnAsnGlnLysGlnAsp-----73
Db      288 CTGGAATCC-----GACAACGAAAAACAGAGAGATTATGTTAAGGAGGGGGCGGAAA 341
Qy      74 -----ThrAspPheSerValLeuLysPhe 81
Db      342 GTGAAGGAGAAATGGAGAAACTAGTAATGGCAACGGAACCTGATGTTATGGCGCTCAAAATTC 401
Qy      82 AlaTyrArgProSerValProAlaHisArgLysValLysGluSerProLeuSerSerAsp 101
Db      402 ACATTTCAGCGCGCGCGCGCTGCTCACGCAAAAATAAGGAGAGTCTCTTAGCTCCGAC 461
Qy      102 ThrIlePheArgGlnSerHisAlaGlyLeuPheAsnLeuCysIleValValLeuValAla 121
Db      462 GCCATCTTCAACACAGAGCATGCGAGGCTCTTCAACCTTTGTATAGTGGTCTGTGCTT 521
Qy      122 ValAsnSerArgLeuIleIleGluAsnLeuMetLysTyrGlyTyrLeuIleLysSerGly 141
Db      522 GTAAATAGCAGCAATAAATTAAGTAATTAAGTATGGTGGCTGATCAAAATCAGGA 581
Qy      142 PheTrpPheSerSerLysSerLeuArgAspTrpProLeuPheMetCysLysSerLeu 161
Db      582 TTTTGGTTAGTTCAACATCGCTTAGGGATTGGCCACTGCTAAATGCTGTGTCTTAGTCTT 641
Qy      162 ValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIlePro 181
Db      642 CCAGTTTTCGACATCGCTTCATTTCTTGTGCGAGAAAGTTGGTGAACATAAATATATACCT 701
Qy      182 GluProValValValLeuHisIleIleIleThrSerThrSerLeuPheTyrProVal 201
Db      702 GAGTGGGTGCGCAGTCTTTCTTCTATGTTACAATCAACACAGTGGAAATCTTGTTCAGTT 761
Qy      202 LeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheSer 221
Db      762 GTTCTCATCTTAGTGTGATCTGCTGTTCTATCAGGTGTACGCTAATCTCTTGTCT 821
Qy      222 CysValValTyrLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetCargAlaLeu 241
Db      822 TGCACCTGTATGTTGAAGTCTGTTCTTCTACCCACATACAAACTATGATTTGAGAGTACTT 881
Qy      242 ThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeuAsnMetAspTyrProTyr 261
Db      882 GCAAAATACACTGTATAGTGGGAAAGTACTTCCAGGTACTGGAACCTCGACTAGCTTAT 941
Qy      262 AsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyrGlnPro 281
Db      942 GATGTAAGCTTTAAGAGTCTGGCATACTTCATGGTGTCTCTACATTTGTGTACAGGCA 1001
Qy      282 SerTyrProArgThrProTyrIleArgLysGlyTyrLeuPheArgGlnLeuValLysLeu 301
Db      1002 AGCTACCTCGGACAGCTTGCATTCGGAAGGGTTGGTGGTAAAGCAACTAATTAAGCTG 1061
Qy      302 IleIlePheThrGlyValMetGlyPheIleIleAspGlnTyrIleAsnProIleValGln 321
Db      1062 GTAAATATTCAGAGACTCATGGGATTTATTATAGAACAGTACATAAACCCTGATCGTTCAA 1121
Qy      322 AsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeu 341
Db      1122 AATTCTCAACATCCTCTGAAGAGGAACCTTTTATGCCATTGAGAGGGTCTTGAAGCTT 1181
Qy      342 SerValProAsnLeuTyrValTyrLeuCysMetPheTyrCysPhePheHisLeuTrpLeu 361
Db      1182 TCTGTTCCAAATTTATATGTTGGCTCTGCGATGTTTATTTGTTTTTCCACCTCTGGCTA 1241
Qy      362 AsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPheTyrLysAspTrpTrpAsn 381
Db      1242 AATATATTCTGTAACATCTCTGCTTTGGGACCGTGGATTTTATAAGGATTTGGTGAAT 1301
Qy      382 AlaLysThrValGluAspTyrTyrArgMetTrpAsnMetProValHisLysTrpMetIle 401
Db      1302 GCGAGGACAGTGGAGGAGTACTGGGAATGTGGGAATATGCTGTGCCATAAATGATGTT 1361

```





Db	682	ACGTTGATGCTCCTCCTACCTTGCAATGTGTGGCGCAAGTGGTTCTTACCTCATCTAAC	741
Qy	236	TyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeuAspThrLeu	255
Db	742	TATGACATAAGAACCCTAGCTAATTCATCTGATAAGCC-----	780
Qy	256	AsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaPro	275
Db	781	AATCTGAGTCTCCTCATGTATGCTTGAAGAGCTTGGCGTATTCATGCTTCC	840
Qy	276	ThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeuPhe	295
Db	841	ACATTTGTTATACCGGAGCTATCCACGTTCTCCATGTATCCGGAAGGTTGGGTGGCT	900
Qy	296	ArgGlnLeuValLysLeuIlePheThrGlyValMetGlyPheIleIleAspGlnTyr	315
Db	901	CCTCAATTTGCAAGCTGATCATATTCCTGCTGATTCATGGGATTTAATAGACGAATAT	960
Qy	316	IleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThr	335
Db	961	ATAAATCTTATTGTAGGAACTCAAAACATCCTTTGAAAGGGATCTCTTATACGGTGT	1020
Qy	336	GluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMetPheTyrCys	355
Db	1021	GAAGAGTGTGAAGCTTTCAGTTCCAAATTTATAGTGTGGCTCTGCATGTTCTACTGC	1080
Qy	356	PhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArgGluPhe	375
Db	1081	TTCTTCCACCTTTGGTTAAACATATTGCGACAGCTCCTCTGCTCGGGGATCGTGAATTC	1140
Qy	376	TyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrpAsnMetPro	395
Db	1141	TACAAGATTTGGTGAATGCAAAAAGCGTGGAGATATTATGGAGAAATGTGGAATATGCT	1200
Qy	396	ValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGlyLeuProLys	415
Db	1201	GTTCAATAATGATGGTTCGACATGTATCTTCCGTGCTCGCAGAAATATACGAAA	1260
Qy	416	AlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeuCysIleAla	435
Db	1261	GTACCGCTATTATCTGCTTCTTAGTCTCTGCAGTCTTTCATGAGTTATGCATCGCA	1320
Qy	436	ValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGlnValProLeu	455
Db	1321	GTTCTTGTGCTGCTTCAAACTATGGGCTTCTTGGGATTTATGTTACAGTGGCTTTG	1380
Qy	456	ValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGlyAsnMetIle	475
Db	1381	GTATTATACAAACTACCTACAGAAAGGTTT---GGCTCCATGGTGGGAAACATGATA	1437
Qy	476	PheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyrTyrHisAsp	495
Db	1438	TTCTGTTTACCTTCTGCAATTTTCGACAACCGATGTGTGCTTCTTTATTATACGAC	1497
Qy	496	LeuMetAsnArgLysGlyLysLeu	503
Db	1498	TTGATGAACCGCAAGGAAGATG	1521
RESULT 4			
LOCUS	AY084052	2090 bp	linear
DEFINITION	Tropaeolum majus putative diacylglycerol acyltransferase mRNA,		
	complete cds.		
ACCESSION	AY084052		
VERSION	AY084052.1		
KEYWORDS	GI:20086323		
SOURCE	Tropaeolum majus.		
ORGANISM	Tropaeolum majus		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	Rosidae; eurosids II; Brassicales; Tropaeolaceae; Tropaeolum.		
REFERENCE	1 (bases 1 to 2090)		

AUTHORS	Mietkiewska, E., Pedersen, K., Katavic, V. and Taylor, D. C.		
TITLE	Characterization of a putative diacylglycerol acyltransferase mRNA from Tropaeolum majus embryo		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2090)		
AUTHORS	Mietkiewska, E., Pedersen, K., Katavic, V. and Taylor, D. C.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-MAR-2002) Seed Oil Biotechnology, National Research Council of Canada, Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada		
FEATURES	Location/Qualifiers		
source	1..2090		
	/organism="Tropaeolum majus"		
	/db_xref="taxon:4020"		
	171..1727		
CDS	/codon_start=1		
	/product="putative diacylglycerol acyltransferase"		
	/protein_id="AA03340.1"		
	/db_xref="GI:20086324"		
	/translation="MAVAESQNTTMSGHGSDLNPNRRKPSSSVIPSSSGFTST NGVPATGHAENRQDRVGAENATGSYNLIGNGGVVIGNEEKQVGTDIRFTYRPS FPAHRRVRESPLSDAIFKQSHAGLNLIVLIVANSRLIENLMKYGLWIDYGFWF NSRSLGWSIFMCCLTLPFFLAAFIVEKLVORHISLVAVLHVIVSTAALYPIV ILVLCDSVYMSVVLMLFGCIMWLKLSYATSSDIPFLAKSGYKGDHPNSTIVSCS YDLSKSLAYEFWAPTLCYQFSYRSCIRKGVVRQFKLIVTIGLMPTIEQITNP IVRSKKPLKGDLYAIEVLKLVSPNLYWLCMPYSFFHLMLILAEELFGDREFY KDNWNAKTVAEYWKWNNPVRHLYFPLLRNGIPKEGAILIAFLVSGAFHELCI AVPCHVFKLWAFIGIMFOVPLVLTINYLQERFSNWSVGNMIFWTFICILGPMCVLLY YHDLINLKEK"		
BASE COUNT	527 a	397 c	458 g 708 t
ORIGIN	US-09-856-018B-16 (1-504) x AY084052 (1-2090)		
Alignment Scores:			
Pred. No.:	4,87e-153	Length:	2090
Score:	1836.00	Matches:	342
Percent Similarity:	78.05%	Conservative:	67
Best Local Similarity:	65.27%	Mismatches:	87
Query Match:	68.51%	Indels:	28
DB:	8	Gaps:	5
Qy	1	MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer	18
Db	171	ATGGCGGTGGCAGATCGTGCAGAACACGACACCATGAGTGGTCACGGCGATCGGAT	230
Qy	19	-----LeuArgArgProSerAla-----	25
Db	231	CTCAACAATTTCCGTAGAGAAACCGAGTTCCTCCGTGATTGAACCTTCGTCGTCGGT	290
Qy	26	---ThrSerThrAlaGlyLeuPheAsnSerProGluThrThrThrAspSerSerGlyAsp	44
Db	291	TTTACATCCACCAATGGGCTA-----CGGGCGATGGCCAGTGGCTGAGAATCGT	341
Qy	45	Asp-----LeuAlaLysAspSerGlySerAspAspSerIleAsnSer	58
Db	342	GACCAAGGATCGGGTATGGGCTATGAGAACGCAACAGGATCGGTCAACTTAATTGGAAT	401
Qy	59	AspAspAlaAlaValAsnSerGlnGlnAsnGluLysGlnAspThrAspPheSerVal	78
Db	402	GGTGAGGCGGTGTTATCGGGAATGAAGAAACAGGATAGGGGAGACTGAT-----	452
Qy	79	LeuLysPheAlaTyrArgProSerValProAlaHisArgLysValLysGluSerProLeu	98
Db	453	ATACGATTCACTTACCGGCTTCGTTTCGGCTCATCGGAGGTGAGGAGAGTCTCTTT	512
Qy	99	SerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPheAsnLeuCysIleValVal	118
Db	513	AGCTCTGATGCAATCTTCAACAGAGCCCATCGGGTTTATTCAACTGTGTAGTAGTG	572
Qy	119	LeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMetLysTyrGlyTrpLeuIle	138
Db	573	CTCATTCGATTAACAGTAGGCTTATCATCGAAATCTTATCAAGTATGCTGGTTGATC	632

Qy	499	ArgLysGlyLys	502
Db	1713	CTAAAGGAAAAG	1724
RESULT	5		
LOCUS	AF051849		
DEFINITION	Arabidopsis thaliana diacylglycerol acyltransferase (DAGAT) mRNA, complete cds.		
ACCESSION	AF051849	AF061563	
VERSION	AF051849.1	GI:6625552	
KEYWORDS	Arabidopsis thaliana.		
SOURCE	Arabidopsis thaliana.		
ORGANISM	Arabidopsis thaliana.		
REFERENCE	1 (bases 1 to 1845)		
AUTHORS	Bouvier-Nave,P., Benveniste,P., Oelkers,P., Sturley,S.L. and Schaller,H.		
TITLE	Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase		
JOURNAL	Eur. J. Biochem. 267 (1), 85-96 (2000)		
MEDLINE	20069349		
PUBMED	10601854		
REFERENCE	2 (bases 1 to 1845)		
AUTHORS	Benveniste,P.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-FEB-1998) Institut de Biologie Moleculaire des Plantes, Centre National de la Recherche Scientifique, 28 rue Goethe, Strasbourg 67083, France		
COMMENT	On Jan 24, 2000 this sequence version replaced gi:6650202.		
FEATURES	Location/Qualifiers		
source	1..1845		
gene	/organism="Arabidopsis thaliana"		
CDS	/cultivar="Columbia"		
	/db_xref="taxon:3702"		
	/clone_lib="ABRC library, catalog number CD4-15"		
	/dev_stage="3 day-old seedling"		
	/note="screened with EST, GenBank Accession Number AA042298"		
	1..1845		
	/gene="DAGAT"		
	88..1650		
	/gene="DAGAT"		
	/function="catalyzes the acylation of diacylglycerol by oleoyl-CoA to give triacylglycerol"		
	/note="acyl-CoA cholesterol acyltransferase homolog"		
	/codon_start=1		
	/product="diacylglycerol acyltransferase"		
	/protein_id="AA19462.1"		
	/db_xref="GI:6625553"		
	/translation="MAILDSAGVTVTENGGEFVLDLRLRRKSRSDSNGLLSSG DNNPSDDVGPADVRDRIDSVVNDDAQGTANLAGNGDNGNGGGGCGRGNAD ATPTYRPSYPARRARSLSDAIFQSHAGLNLVAVVLLAVNSRLIENLMKYGM LRTDWFSSRLRDWPLEWCISLSEFLAAFTVEKLQVISPVPVIFLHIITM TEVLYPVYTLRCDSAFLSGTQMLLTCIWLKLVSAFTSDIRSLANAADKANPEV SYVLSKLAYFMVAPTLCQPSYRSACIRKGVARFAKLVIFFTFMGFIETQIRN PIVRNSKHLKGLDLYAIERVLKSLVPLNYVLCMFYCFHLMNLIAELLCFGEFR YKDMNNAKSVGDYWRMNMVPHKWMVRHYVPCLRSKIPKTLAIILAFILVSAVHELC IAVPCRLFKLWAFGLGIMFQVPLVFITNLOERFGSVGNKIFWFIICFQGPCMVCVLY YHDLNMRKGSMS"		
BASE COUNT	453 a	378 c	421 g 593 t
ORIGIN			
Alignment Scores:			
Pred. No.:	8,93e-152	Length:	1845
Score:	1821.00	Matches:	353
Percent Similarity:	75.33%	Conservative:	47
Best Local Similarity:	66.48%	Mismatches:	91
Query Match:	67.95%	Indels:	40
DB:	8	Gaps:	9





/translation="MAILDSAGVTTTENGGEFVDLDRRLRRKSRSDSNGLLLSGS  
DNPSDDVGPADVDRSDSVNDDAQGTANLAGDNGGDNNGRGGGGRGNAD  
ATFTYRPSVPAHRRRSLSDAIFKQSHAGFLNLCVVLIAVNIENLMKYGW  
LIRTFDFWSSRSLRDMPCCLSLSTFPLAFTVVKLVQKYSPPVWIFLHIITM  
TELYVPVYTLRCDSPFLSGVTLMLTICVWLKLVSAHTSYDIRISLANAADKANPEV  
SYVSLKSLAYENVATLQCPSPRSACIRKQWARQAKLVIFTFGFWGFIIEQIN  
PIVRNKKPLKGLDLYATIERVLKLSVPLNLYVWICMFCFFHMLNLAELLCFGDRF  
YKDWNAKSGVDYWRMNPVHKWYHIFPCIRSKIPKTLAIILAFIVSAVFHELC  
IAVPCRLFLKWLAFGLMFOVPLVFNLYQERFSGTVGNMFIFCIFQPMCVLLY  
YHDLNMRKSGMS"

BASE COUNT 458 a 399 c 431 g 616 t  
ORIGIN

## Alignment Scores:

Pred. No.: 9, 27e-152 Length: 1904  
Score: 1821.00 Matches: 353  
Percent Similarity: 75.33% Conservative: 47  
Best Local Similarity: 66.48% Mismatches: 91  
Query Match: 67.95% Indels: 40  
DB: 8 Gaps: 9

US-09-856-018b-16 (1-504) x ATH238008 (1-1904)

Qy 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18  
Db 139 ATGCGGATTTGGATTCT---GCTGGCGTTACTACGGTGACGGAGAACGGTGGCGGAGAG 195  
Qy 19 -----LeuArgArgArgProSer---AlaThrSerThrAlaGlyLeu 31  
Db 196 TTCGTCGATCTGATAGCTTCGTCGAGGAAATCGAGATCGGATCTTCTACCGGACTT 255  
Qy 32 PheAsnSerProGluThrThrAspSerSerGlyAspLeuAlaLysAspSerGly 51  
Db 256 CTTCTCTCTGTTCCGATAATAATCTCTCTCGGATGATGTTGGAGCTCCCGCGGACGTT 315  
Qy 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67  
Db 316 AGGATCGGATGATTCGGTGTAAAGATGACGCTCAGGAGACAGCCAAATTTGGCCGGA 375  
Qy 68 GlnAsnGlu-----LysGln 72  
Db 376 GATAATAACGGTGGTGGCGATAATAACGGTGGTGGAGAGCGCGGAGAGGAGGA 435  
Qy 73 AspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHisArgLys 92  
Db 436 AACCGCGATGCTAGC-----TTACGATACGACCGTCGGTTCACGCTCATCGGAGG 486  
Qy 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112  
Db 487 GCGAGAGAGTCCACTTACTCGGACGCAATCTTCAACAGAGCCATGCCGGATTATTC 546  
Qy 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleLeuGluAsnLeuMet 132  
Db 547 AACCTCTGTGTAGTAGTCTTATTGCTGTAAACAGTAGACTCATCATCGAAAAATCTTATG 606  
Qy 133 LysTyrGlyTrpLeuIleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrp 152  
Db 607 AAGTATGTTGGTGTAGTACAGACGGATTTCTGTTTATAGTTCAAGATCGCTCGGAGATTG 666  
Qy 153 ProLeuPheMetCysCysLeuSerValValPheProPheAlaAlaPheIleValGlu 172  
Db 567 CCGCTTTTCATGTGTGTATATCCCTTCGATCTTCTTGGCTGGCTTACGGTTGAG 726  
Qy 173 LysLeuAlaGlnArgLysCysIleProGluProValValValLeuHisIleIle 192  
Db 727 AAATTGGTACTTCAGAAATACATATCAGAACCTGTGTCTATCTTCTTCATATTATTC 786  
Qy 193 ThrSerThrSerLeuPheThrProValLeuValIleLeuArgCysAspSerAlaPheVal 212  
Db 787 ACCATGACAGAGGTTTGTGTACAGTTACCTCACCCCTAACGGTGTGATCTGCTTTT 846  
Qy 213 SerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTala 232

Db 847 TCAGGTCTCACTTTGATGCTCCTCCTACCTTGCATTGTGGCTAAAGTTGGTTCTTTATGCT 906  
Qy 233 HisThrAsnTyrAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252  
Db 907 CATACTAGCTATGACATAGATCCCTAGCCAAATCCAGCTGATAAGGCC----- 954  
Qy 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTatPheLeu 272  
Db 955 -----AATCCTGAAGTCTCCTACTACCTTACCTTGAAGAGCTTGGCATATTTTCATG 1005  
Qy 273 ValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyrIleLeuGlyGly 292  
Db 1006 GTGCTCCCATTTGTTTATTCAGCCAAATTCATCCAGCTTCTGCATGTATACGGAAGGT 1065  
Qy 293 TrpLeuPheArgGlnLeuValLysLeuIlePheThrGlyValMetGlyPheIleIle 312  
Db 1066 TGGTGGCTGCTCAATTTGCAAACTGCTCATATTCACCGATTCATGGGATTTATATA 1125  
Qy 313 AspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332  
Db 1126 GAACAATATATAATCTATTGTCAGGAACCTCAAGACATCCTTTGAAAGCGCATCTTCTA 1185  
Qy 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352  
Db 1186 TATGCTATTGAAAGAGTGTGAAGCTTTCAGTTCCTCAAAATTTATATGTGGCTCTGCATG 1245  
Qy 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372  
Db 1246 TTCTACTGCTTCTCCACCTTTGGTAAACATATTTGCGACAGCTTCTCTCGTTCGGGAT 1305  
Qy 373 ArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392  
Db 1306 CGTGAATCTCAAAAGATTTGGTGAATGCAAAAGTGTGGAGATTACTGGAGAACTGG 1365  
Qy 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412  
Db 1366 AATATGCTGTTCAATAATGATGTTGACATATATATCTTCCGCTGTTGCGGCAAG 1425  
Qy 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432  
Db 1426 ATACAAAGACACTCGCCATTATCATTTCTTCTACTCTCTGCAGTCTTTCATGACCTA 1485  
Qy 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452  
Db 1486 TGCATCGACCTTCTTCTGCTCTCTTCAAGCTATGGGCTTTCTTGGGATTTATGTTTTCAG 1545  
Qy 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472  
Db 1546 GTGCTTTGGTCTTTCATCACAATACTATCTACAGGAAAGGTTT---GGCTCAACGGTGGG 1602  
Qy 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlnProMetCysValLeuLeuTyr 492  
Db 1603 AACATGATCTTCTGTTTCATCTTCTGCAATTTTCGGAACACCGATGTGTGCTTCTTTAT 1662  
Qy 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503  
Db 1663 TACCACGCTGATGAACGAAAGGATCGATG 1695

## RESULT 7

LOCUS AX090345 1942 bp DNA linear PAT 21-MAR-2001  
DEFINITION Sequence 38 from Patent WO0116308.  
ACCESSION AX090345  
VERSION AX090345.1 GI:13444207  
KEYWORDS  
SOURCE thale cress  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 1942)  
REFERENCE Lasserre M. and van Eenennaam A.  
AUTHORS Plant sterol acyltransferases  
TITLE



JOURNAL Patent: WO 0116308-A 38 08-MAR-2001;  
 MONSANTO COMPANY (US)  
 FEATURES Location/Qualifiers  
 source 1..1942  
 /organism="Arabidopsis thaliana"  
 /db\_xref="taxon:3702"  
 BASE COUNT 461 a 421 c 425 g 635 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 9.5e-152 Length: 1942  
 Score: 1821.00 Matches: 353  
 Percent Similarity: 75.33% Conservative: 47  
 Best Local Similarity: 66.48% Mismatches: 91  
 Query Match: 67.95% Indels: 40  
 DB: Gaps: 9

US-09-856-018B-16 (1-504) x AX090345 (1-1942)

```

QY 1 MetaLal1eSerAspGluProGluSerValAlaThraLeuAlaSnHisSerSer----- 18
   ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 237 ATGGCGATTGTGATTCT--GCTGCGGTTACTACGGTCACGGAGACGGTGGCGGAGAG 293

QY 19 -----LeuArgArgArgProSer--AlaThrSerThraLeuGlyLeu 31
   ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 TTCGTGCATCTGATAGGCTTCGTGCACGGAATCGAGATCGGATCTTCTTAACGGACTT 353

QY 32 PheAsnSerProGluThrThrAspSerSerGlyAspLeuAlaLysAspSerGly 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 CTTCCTCTCTGGTCCGATAAATTAATCTCCTCGGATGATGTTGGAGCTCCGCCGACGTT 413

QY 52 SerAsp-----AspSer1eAsnSerAspAspAla-----AlaValAsnSerGlnGln 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 414 AGGATCGGATTGATTCGGTTGTTAACGATGACGCTCAGGGAACAGCCAAATTGGCCGGA 473

QY 68 GlnAsnGlu-----LysGln 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 474 GATAAATAACGGTGGTGGCGGATAATAACGGTGGTGGAAAGCGCGCGGAAGAAGAGGA 533

QY 73 AspThrAspPheSerValLeuLysPheAlaTyArgProSerValProAlaHisArgLys 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 534 AACCCGATGCTACG-----TTTACGTATCACCACGCTCGGTTCCAGCTCATCGGAG 584

QY 93 ValLysGluSerProLeuSerSerAspThrLlePheArgGlnSerHisAlaGlyLeuPhe 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 585 GCGAGAGAGAGTCCACTTAGTCCGACGCAATCTTCAACAGAGCCATGCCGGATTATTC 644

QY 113 AsnLeuCysLleValValLeuValAlaValAsnSerArgLeuLleGluAsnLeuMet 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 645 AACCTCTGTAGTAGTCTTATTGCTGTAAACACATAGACTCATCATCGAAATCTTATG 704

QY 133 LysTyrgLlyrPlrLeuLleLysSerGlyPheTrpPheSerSerLysSerLeuArgAspTrp 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 705 AAGTATGGTGGTGTATCAGACGAGATTCTCGTGTAGTTCAAGATCGCTCGAGATGG 764

QY 153 ProLeuPheMetCysLysLeuSerLeuValValPhePropheAlaAlaPheLleValGlu 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 765 CCGCTTTTCATGTTGTATATCCCTTTTCGATCTTCTTTCGCTGCCTTTACGGTTGAG 824

QY 173 LysLeuAlaGlnArgLysCysLleProGluProValValValLeuHisLleLle 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 825 AAATTGGTACTTCAGAAATACATATACAGAACCTGTGTCTTCTTCATATTATTATC 884

QY 193 ThrSerThrSerLeuPheTyProValLeuValLleLeuArgCysAspSerAlaPheVal 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 885 ACCATCAGCAGAGGTTGTGTACGATTACGTCACCCCTAAAGGTGTGATCTGCTTTTATA 944

QY 213 SerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSerTyrAla 232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 945 TCAGGTGTCACTTTGATGCTCTCATTGCATTTGCTGGCTAAAGTTGCTTCTTATGCT 1004

QY 233 HisThrAsnTyAspMetArgAlaThrLysLeuValGluLysGlyGluAlaLeuLeu 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



Oilseeds Research, John Innes Centre, Colney Lane, Norwich, NR4 7UH, U.K

FEATURES  
source Location/Qualifiers  
1..1988 /organism="Arabidopsis thaliana"  
/sub\_species="Columbia"  
/db\_xref="taxon:3702"  
231..1793 CDS  
/EC\_number="2.3.1.20"  
/codon\_start=1  
/product="diacylglycerol O-acyltransferase"  
/protein\_id="CAB44774.1"  
/db\_xref="GI:5050913"  
/translation="MALDSAGVTVTENGGEFVDLRLRRKRSRSDSSNGLLLSGS  
DNPSDDVADPADVRIDSVVNDDAQTNLAGDNNGGDNNGGGGGGGGEGGNAD  
ATFYRSPVARRRAREPLSDAIFKQSHAGLFNLCVVLIAYNSRIIENLMKYGW  
LIRDFVSSRLSDWPLMCCIISLIFPLAAFTVKILQKYESEPVIFLHIITM  
TEVLYPVVTURCDASFLSGVTLMLTCTIVMLKLVSAHTSVDRSLANAADKANPEV  
SYVYSLKSLAYFMVAPTLCYOPSPRSACIRKGMVAROFKLVIFGTGPMGFLIBQYIN  
PIVNSKHPLAGDLLIALERVLKLSVPLNYVLMCMFYCFPHLNLIAELLCLFGDREF  
YKDMNNAKSGVDIWRMMNMPVHKWVRHIYFPCLSKIPKTLALIIILVSAVFELC  
IATVCRFLKFWAFIGMFQVPLVITNYLQERFSGTVGNMIFWIFCFIFGQPMCVLLY  
YHDLNRRKGSN"

BASE COUNT 478 a 423 c 436 g 651 t

ORIGIN

Alignment Scores:  
Pred. No.: 9,77e-152 Length: 1988  
Score: 1821.00 Matches: 353  
Percent Similarity: 75.33% Conservative: 47  
Best Local Similarity: 66.48% Mismatches: 91  
Query Match: 67.95% Indels: 40  
DB: 8 Gaps: 9

US-09-856-018B-16 (1-504) x ATH131831 (1-1988)

Qy 1 MetAlaIleSerAspGluProGluSerValAlaThrAlaLeuAsnHisSerSer----- 18  
Db 231 ATGGCGATTGCGATCTCTCTCGACGGAATCGAGATCGGAGCGGAGAG 287  
Qy 19 -----LeuArgArgProSer-----AlaThrSerThrAlaGlyLeu 31  
Db 288 TTCGTCGATCTTGTATAGGCTTCGTCAGCGAATCGAGATCGGATCTCTTAACGGACIT 347  
Qy 32 PheAsnSerProGluThrThrThrAspSerGlyAspAspLeuAlaLysAspSerGly 51  
Db 348 CTTCCTCTCGGTTCGGATAATAATCTCTCTCGGATGATGTTGGAGCTCCCGCGACGTT 407  
Qy 52 SerAsp-----AspSerIleAsnSerAspAla-----AlaValAsnSerGlnGln 67  
Db 408 AGGGATCGGATGATTCGCTGTTAACGATGACGCTCAGGGAACAGCCCAATTTGGCCGGA 467  
Qy 68 GlnAsnGlu-----LysGln 72  
Db 468 GATAATAACGGTGGTGGCCATAATAACGGTGGTGGAGAGCGCGGAGAGAGAGA 527  
Qy 73 AspThrAspPheSerValLeuLysPheAlaTyArgProSerValProAlaHisArgLys 92  
Db 528 AACGGCGATGCTACG-----TTTACGATACGACGCTCGGTTCCAGCTCATCGGAGG 578  
Qy 93 ValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGlyLeuPhe 112  
Db 579 GCGAGAGAGATGCTTACGTCGCGGCAATCTTCAACAGAGCGCATCGCGGATTTATTC 638  
Qy 113 AsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsnLeuMet 132  
Db 639 AACCTCTGTAGTAGTCTTATTCTGTAAACAGTAGACTCATCATCGAAAATCTTAGT 698  
Qy 133 LysTyGlyTrpLeuIleLysSerGlyPheThrPheSerSerLysSerLeuArgAspTrp 152  
Db 699 AAGTATGGTGTGATCAGAACGGATTTCTGTTTAGTTCAAGATCGTTCGAGATTGG 758  
Qy 153 ProLeuPheMetCysCysLeuSerLeuValValPheProPheAlaAlaPheIleValGlu 172

Db 759 CCGCTTTTCATGTTGTTATATCCCTTTCCATGCTTTCCCTTTGGCTGCCCTTTACGGTTGAG 818  
Qy 173 LysLeuAlaGlnArgLysCysIleProGluProValValValValValLeuHisIleIleIle 192  
Db 819 AAATGGTACTTTCAGAAATACATATCAGAACCTGTTGTCATCTTCTTCATATATTATTC 878  
Qy 193 ThrSerThrSerLeuPheThrProValLeuValIleLeuArgCysAspSerAlaPheVal 212  
Db 879 ACCATGACAGAGGTTTGTATCCAGTTTACCTACCTCAGCTTAAAGTTGGTTCTTATGCT 938  
Qy 213 SerGlyValThrLeuMetLeuPheSerCysValValThrLeuLysLeuValSerTyrAla 232  
Db 939 TCAGGTGTCACCTTGTATGCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 998  
Qy 233 HisThrAsnTyAspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuLeu 252  
Db 999 CATACTAGTATGACATAAGATCCCTAGCAATGACGCTGATAAGGCC----- 1046  
Qy 253 AspThrLeuAsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeu 272  
Db 1047 -----AATCCTGAAGTCTCTACTACTAGTCTTGAAGAGCTTGCATATTTTCATG 1097  
Qy 273 ValAlaProThrLeuCysTyrGlnProSerTyrProAspThrProTyrIleArgLysGly 292  
Db 1098 GTCCGCTCCACATTTGTTTATCAGCCAAAGTTATCCACGTTCTGCATGTATACGAAGGT 1157  
Qy 293 TrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIle 312  
Db 1158 TGGGTGGCTCGTCAATTTGCCAAACTGGTCATATTCACCGGATTCATGGGATTTATAATA 1217  
Qy 313 AspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332  
Db 1218 GAACAATATATAATCTTATTCAGGAATCTTGAAGAGCTTGAAGAGCTTCTCTA 1277  
Qy 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352  
Db 1278 TATGCTATTGAAAGAGTGTGAAGCTTTCAGTTCCAAATTTATATGTGTGGTCTGTCATG 1337  
Qy 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372  
Db 1338 TTCTACTGCTTCTCCACCTTGGTTAAACATATTTGGCAGAGCTTCTCTGCTTCGGGAT 1397  
Qy 373 ArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392  
Db 1398 CGTGAATTTACAAAAGATTGGTGAATGCAAAAGTGTGGAGATTACTGGAGATGTGG 1457  
Qy 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412  
Db 1458 AATATGCCCTGTTCAAAATGGATGGTTCGACATATATATCTCCCGTGTTCGCGAGCAAG 1517  
Qy 413 LeuProLysAlaAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432  
Db 1518 ATACCAAGACACTGCCCATTTATCTGTTTCTTAGTCTCTGCAGTCTTTCATGAGCTA 1577  
Qy 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyIleMetPheGln 452  
Db 1578 TGATCGCAGTTCCTTGTGCTCTCTCAAGCTATGGCTTTCTTGGGATTTATGTTTCAG 1637  
Qy 453 ValProLeuValIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472  
Db 1638 GTGCCCTTGGTCTTCATCAAACTATCTACAGAAAGGTTT---GGCTCAACGGTGGG 1694  
Qy 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492  
Db 1695 AACATGATCTTCTGGTTCATCTTCTGCAATTTTCGGACAAACCGATGTGTGTCTTTAT 1754  
Qy 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503  
Db 1755 TACCACGACCTGATGAACCGAAAGGATCGATG 1787  
RESULT 10  
AT05480



```

QY 313 AspGlnTyrIleAsnProfilValGlnAsnSerGlnHisProLeuLysGlyAsnLeuLeu 332
Db 1149 GAACAATATATAAATCTTATCTCAGGAACCTCAAGACACTCTTTGAAGCGCATCTTCTTA 1208
QY 333 TyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCysMet 352
Db 1209 TATGCTATTGAAGAGTGTGAAGCTTTCAGTTCCTCCAAATTTATATGTGTGGCTCTGCATG 1268
QY 353 PheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeuArgPheGlyAsp 372
Db 1269 TTCTACTGCTTCTTCACCTTTGGTAAACATATTGGCAGAGCTTCTCTGCTTCGGGAT 1328
QY 373 ArgGluPheTyrLysAspTrpAsnAlaLysThrValGluAspTyrTrpArgMetTrp 392
Db 1329 CQTGAATCTCACAAGATTGGTGGAAATGCAAAAGTGTGGGAGATTACGGAGAATGTGG 1388
QY 393 AsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHisGly 412
Db 1389 AATATGCTGTTTCATAAATGGATGGTTCGACATATATACTTCCCGTGTGGCAGCAG 1448
QY 413 LeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGluLeu 432
Db 1449 ATACAAAGACACTCGCCATTATCATGTTCTTCTAGTCTCTGCACTCTTTCATGAGTCA 1508
QY 433 CysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPheGln 452
Db 1509 TGCATCGCAGTTCCTGTGCTCTCTCAAGCTATGGCTTTCTTGGGATTATGTTTCAG 1568
QY 453 ValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetValGly 472
Db 1569 GNGCCTTTGGTCTTCATCACAACTATCTACAGGAAGGTTT--GGCTCAACGGTGGG 1625
QY 473 AsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyr 492
Db 1626 AACATGATCTTCTGTTTCATCTCTCGCAATTTTCGGACACCGATGTGTGCTTCTTTAT 1685
QY 493 TyrHisAspLeuMetAsnArgLysGlyLysLeu 503
Db 1686 TACCAGCCTGTGATGACCGAAAGATGTCATG 1718

RESULT 11
AF164434 1512 bp mRNA linear PLN 30-NOV-1999
LOCUS Brassica napus putative diacylglycerol acyltransferase (DGAT1)
DEFINITION mRNA, complete cds.
ACCESSION AF164434
VERSION AF164434.1 GI:5579407
KEYWORDS
SOURCE Brassica napus.
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
AUTHORS Nykiforuk,C.L., Laroche,A. and Weselake,R.J.
TITLE Isolation and Characterization of a cDNA Encoding a Second Putative
Diacylglycerol Acyltransferase from a Microspore-derived Cell
Suspension Culture of Brassica napus L. cv Jet Neuf (Accession No.
AF164434). (PCR99-158)
JOURNAL Plant Physiol. 121 (3), 1053 (1999)
REFERENCE
AUTHORS Nykiforuk,C.L., Laroche,A. and Weselake,R.J.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1999) Chemistry and Biochemistry, University of
Lethbridge, 4401 University Drive, Lethbridge, Alberta T1K 3M4,
Canada
FEATURES
Source Location/Qualifiers
1..1512
/organism="Brassica napus"
/strain="cultivar Jet Neuf"
/db_xref="taxon:3708"
/note="microspore-derived cell suspension culture"
1..1512

```

## CDS

```

/gene="DGAT1"
1..1512
/EC_number="2.3.1.20"
/codon_start=1
/product="putative diacylglycerol acyltransferase"
/protein_id="RAD45536.1"
/db_xref="GI:5579408"
/translation="MAILDSGGVAVPPTENGVAIDLRLHRRKSSSDSSNGLLSDTSPS
DDVGAARERDRVDSAAEEAEAGTANLAGDAETRESAGGVRFRTYRPSVPAHRRTRE
SPLSDAIFKQSHAGLFNVLAVANSRLIENLMKYGLWIRTFDFWFSSTSLRDWP
LFMCCLSLSPFLAAFTVEKVLQKFISEPVAILUHVITTEVLVYPVTVLRCDSEF
LSGYTLLLTCTIVWLKLVSYAHTSDIRTLANSADKVDPEISYVSLSKSLAFWVAPT
LCYQSPSPRNLYVWLCMFYCFHLLNLIAELLCFGDREYFKDMNNAKSGDYWRMW
IERVLKMWVRHVYFCLRIKIPKPAIIIAFLVSAVFEHLCTIAPVCRFLNFWAPGIM
NPQPLVFTTNFLQERFGSMVGNMIFGASCFIQPMCGLLLYVHDLNRKGSMS"
BASE COUNT 344 a 343 c 363 g 462 t
ORIGIN
Alignment Scores:
Pred. No.: 1..56e-149 Length: 1512
Score: 1794.50 Matches: 335
Percent Similarity: 79.11% Conservative: 55
Best Local Similarity: 67.95% Mismatches: 78
Query Match: 66.96% Indels: 25
DB: 8 Gaps: 5
US-09-856-018B-16 (1-504) x AF164434 (1-1512)
QY 20 ArgArgArgProSerAlaThrSerThrAlaGlyLeuPheAsnSerProGluThrThrThr 39
Db 76 CQTCGTAATTCAGTTCGGATTCTTCCAAACGGACTCTCTCCGAT----- 120
QY 40 AspSerSerGlyAspLeuAlaLysAspSerGlySerAspAspSerIleAsnSerAsp 59
Db 121 ACTTCCCGCTCGACGATGTTGGAGCTGCGCGCGCGGCGAAGGATCGGGTT----- 171
QY 60 AspAlaAlaValAsnSerGlnGlnAsn----- 70
Db 172 GATTCGCTGCGGAGGAGGAGCTCAGGACAGCGCAATTTAGTCGCGGAGATGCCGAA 231
QY 71 LysGlnAspThrAspPheSerValLeuLysPheAlaTyrArgProSerValProAlaHis 90
Db 232 ACTAGGAATCCCGCGGAGCGCATGTAAGTTTACGTATCGACCGCTGGTTCAGTCTAT 291
QY 91 ArgLysValLysGluSerProLeuSerSerAspThrIlePheArgGlnSerHisAlaGly 110
Db 292 CGGAGGACGAGGAGAGTCTCTCAGCTCCGACGCTATCTTCAACAAAGCCATGCAGGA 351
QY 111 LeuPheAsnLeuCysIleValValLeuValAlaValAsnSerArgLeuIleIleGluAsn 130
Db 352 TTGTTCAACCTCTGTGTAGTTCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 411
QY 131 LeuMetLysTyrGlyTrpLeuLysSerGlyPheTrpPheSerSerLysSerLeuArg 150
Db 412 CTCATGAAGTATGGTGGTGTATCAGAACTGATTTTGTGTGTGTGTGTGTGTGTGTGTGT 471
QY 151 AspTrpProLeuPheMetCysLysLeuSerLeuValValPheProPheAlaLapHelle 170
Db 472 GACTGGCGCTTTTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 531
QY 171 ValGluLysLeuAlaGlnArgLysCysIleProGluProValValValValLeuHisIle 190
Db 532 GTCGAGAAATGGTACTTACAGAAATTCATATCTAGAGCTGTGTGCCATCATCTTCTCATGTC 591
QY 191 IleIleThrSerThrSerLeuPheTyrProValLeuValIleLeuArgCysAspSerAla 210
Db 592 ATTATAACCATCAGAGAGTCTTGTATCCAGTCTACGTACACTGAGGTGTGATTCTGCC 651
QY 211 PheValSerGlyValThrLeuMetLeuPheSerCysValValTrpLeuLysLeuValSer 230
Db 652 TTCTGTGTCAGGTGTACGTTGATGTGCTCAGTTCGATGTGCTGAAGTGTGTTTCTTCT 711

```



```
Qy 269 AlaTyrPheLeuValAlaProThrLeuCysTyrGlnProSerTyrProArgThrProTyr 288
Db 400 GCTATTTCATGCTGCTCCACATTTGTTATFACCCAGCATATCCAGCTTCTCCATGTT 459
Qy 289 IleArgLysGlyTrpLeuPheArgGlnLeuValLysLeuIleIlePheThrGlyValMet 308
Db 460 ATCCGGAAGGTTGGTGCTCGTCAATTTGCAAAACTGGTCATATTCAGTGCATCATG 519
Qy 309 GlyPheIleIleAspGlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLys 328
Db 520 GGATTATTAATAGACCAATATATAATCCTATTGTTAGAACTCAAGCATCCTCTGA 579
Qy 329 GlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeuSerValProAsnLeuTyrVal 348
Db 580 GGGGACCTTCTATAGCTATTGAAGAGTGTTGAAGCTTTCAGTCCCAATCTATATGTG 639
Qy 349 TrpLeuCysMetPheTyrCysPheHisLeuTrpLeuAsnIleLeuAlaGluLeuLeu 368
Db 640 TGGCTCTGCATGTTCTACTGCTTCTCCACCTTGGTTAAACATATTGGCAGAGCTCCTC 699
Qy 369 ArgPheGlyAspArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyr 388
Db 700 TCGTTCGGGGACCGTGAATTTACAAAGATTGCTGGAATGCAAAAGCGTTGGAGATTAT 759
Qy 389 TrpArgMetTrpAsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCys 408
Db 760 TGGAGAAATGGAATGCTGCTTCAAAATGATGCTTGGACATGATATACCTTTCCTGCG 819
Qy 409 LeuArgHisGlyLeuProLysAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeu 428
Db 820 CTGGCATCAAGTACCAAAAGTACCGGCATATATCATGCTTCTTAGCTCTGCGAGT 879
Qy 429 PheHisGluLeuCysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGly 448
Db 880 TTTCATGAGTTATGATCCAGTCTCTTCCGCTCTCTTCAATCTATGGCTTTCATGGGA 939
Qy 449 IleMetPheGlnValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsn 468
Db 940 ATATGTTTCAGTCCCTTGGTCTTTATCACAAACTTTTACAGAAAGGTTT---GGC 996
Qy 469 SerMetValGlyAsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCys 488
Db 997 TCCATGGTGGAAACATGATCTTTGGTTCAGCTCTCTTGCATTTTCGGACAACCGATGCT 1056
Qy 489 ValLeuLeuTyrTyrHisAspLeuMetAsnArgLysGlyLysLeu 503
Db 1057 GGGCTTCTTTATACCATGACCTGATGAACCGCAAGGATCCATG 1101
RESULT 13
AC005917
LOCUS AC005917 92822 bp DNA linear PLN 27-FEB-2002
DEFINITION Arabidopsis thaliana chromosome 2 clone F3P11 map CIC06E08,
complete sequence.
ACCESSION AC005917
VERSION AC005917.3 GI:20197478
KEYWORDS HTG
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 92822)
REFERENCE
AUTHORS Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Shen,M., VanAken,S.E.,
Barnstead,M.E., Mason,T.M., Bowman,C.L., Ronning,C.M.,
Benito,M.-I., Carrera,A.J., Creasy,T.H., Buell,C.R., Town,C.D.,
Nierman,W.C., Fraser,C.M. and Venter,J.C.
JOURNAL Unpublished
REFERENCE
AUTHORS Lin,X.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
REFERENCE
AUTHORS Lin,X.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org
On Apr 18, 2002 this sequence replaced gi:6598497.
FEATURES
Location/Qualifiers
1..92822
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/map="CIC06E08"
/clone="F3P11"
1..27852
/note="overlap with BAC clone F27F23
(AC003058:69644..97495)."
```

misc\_feature  
repeat\_region  
repeat\_region  
repeat\_region  
gene  
mrna  
CDS  
repeat\_region  
repeat\_region  
repeat\_region  
gene  
mrna  
CDS  
gene  
mrna  
CDS  
gene  
mrna  
CDS

Town,C.D. and Kaul,S.  
Direct Submission  
Submitted (27-FEB-2002) The Institute for Genomic Research, 9712  
Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org  
On Apr 18, 2002 this sequence replaced gi:6598497.  
Location/Qualifiers  
1..92822  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/map="CIC06E08"  
/clone="F3P11"  
1..27852  
/note="overlap with BAC clone F27F23  
(AC003058:69644..97495)."  
complement(2151..2171)  
/rpt\_family="AT\_rich"  
complement(5234..5268)  
/rpt\_family="AT\_rich"  
5488..5517  
/rpt\_family="AT\_rich"  
6440..8622  
/gene="At2g19420"  
/note="F3P11.2; predicted by genscan"  
join(<6440..6458,6961..7071,8318..>8622)  
/gene="At2g19420"  
join(6440..6458,6961..7071,8318..8622)  
/gene="At2g19420"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="AADI0141.1"  
/db\_xref="GI:4191772"  
/translation="MHEKVPKRTYDSLYNGYLAIFYGDLRYFTVKIENDICVSLV  
WNTGKTSRGRKIRINVRPKIAPAKNKKPKTKEOIMDPDESDVLTSLGFDG  
GIIPLPKSLMLINLPDSFRDELQTMHVSFYLREVGLMEF"  
complement(7164..7241)  
/rpt\_family="(TA)n"  
complement(9348..9369)  
/rpt\_family="AT\_rich"  
9624..9757  
/rpt\_family="AT\_rich"  
9956..12479  
/gene="At2g19430"  
/note="F3P11.3; contains a Trp Asp (WD 40) repeat  
signature"  
join(<9956..10281,10369..10445,10527..10572,10700..10760,  
10849..10902,10993..11073,11165..11204,11477..11577,  
11750..11860,12390..>12479)  
/gene="At2g19430"  
join(9956..10281,10369..10445,10527..10572,10700..10760,  
10849..10902,10993..11073,11165..11204,11477..11577,  
11750..11860,12390..12479)  
/gene="At2g19430"  
/codon\_start=1  
/product="putative WD-40 repeat protein"  
/protein\_id="AADI0142.1"  
/db\_xref="GI:4191773"  
/translation="MYGATNNWDEYRESILKREIETRTVFTAWAPPARISNPDA  
FVWSDCTLAFHSLNLSVSQSGFYSGKQDVMVAEPERVVRAHKGPAVDYKFGED  
EDALLSGDGRVGRWREAFESDSVLSHLKHNHLKLELLELINFQHKGPWGLSPMP  
ENAMSDVPQSGSVTAAGDSACAYCWDVESGRKMTFKGHSYDLHTVTVSSASQILT  
GSEDTARIDWCKTGKCVIGSQDKRSRLRSKMSALDGSLSWLCVGGOKNLALNLP  
ASECVQITPIPAHVQDVMFDEKQVAVGVGGYGVIVDSIQFSGHLCITFRSSSL"  
complement(12903..14545)  
/gene="At2g19440"  
/note="F3P11.4"  
complement(join(<12903..13516,13584..14083,14223..>14545))  
/gene="At2g19440"  
complement(join(12903..13516,13584..14083,14223..14545))  
/gene="At2g19440"  
/codon\_start=1  
/product="putative beta-1,3-glucanase"







QY	295	heArgGlnLeuVallysLeuIleIlePheThrGlyValMetGlyPheIleIleAsp----	313
DB	22722	CTCGTCAATTTTGCAAAACTGGTCATATTACACGGATTATGAGGATTTATAATAGAACAAAG	22781
QY	313	-----	313
DB	22782	TACGTTTTTCACATCTTGCTTTATTAGTTTTCCTTGGTGAATAATCATCATCCCTCGGTGCT	22841
QY	314	-----GlnTyrIleAsnProIleValGln	321
DB	22842	CACCACTTGACTTCATGTTCTTTTGTGTACATTTTGGCAGTATATAAATCTATTATGFCAGG	22901
QY	322	AsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeu	341
DB	22902	AACTCANAAGCATCCTTTGAAGCGCATCTCTATATGCTATTGAAGAGTGTGTGAAGCTT	22961
QY	342	SerValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrp---	360
DB	22962	TCAGTTCCAAATTTATATGTGTGGCTCTGCATGTTCTACTGCTTCTCCACCTTTGGTAT	23021
QY	360	-----	360
DB	23022	GCTGTGATCCCATCTCTTTTCAAAATAAATTGCAAAATTCGAAAAACCGAAAAAGGCTAAAT	23081
QY	360	-----	360
DB	23082	CTCATACGAATTTGATATTTTGTAGTTCTTAGAGTCGGTGTATGTAATTTACGTTACTGAA	23141
QY	361	-----LeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArg	373
DB	23142	CCCAATCTCTTGTCCAAAGGTTAAACATATTTGGCAGAGCTTCTCTGCTTCGGGATCGT	23201
QY	374	GluPheTyrLysAspTrpAsnAlaLysThrValGlu-	386
DB	23202	GAATCTACAAAGATTGGTGGAAATGCAAAAGATGTGGG-AGATGTGAGCATATTTTACTCA	23260
QY	386	-----	386
DB	23261	AAAGAAAACTTATGATTTTAAATGTGTGCTGTTTGGGTGTCATCACTAACTAACCAAAATT	23320
QY	387	-----AspTyrTrpArgMetTrpAsnMet----	394
DB	23321	CATGTATTACACTGCTTTCCTTTATCAGTACTCGAAGATGTGAATATGATGTTGTTCTCT	23380
QY	394	-----	394
DB	23381	TCCTAAACATCACCTTCITTTCTACACAAATAGAGAAGAGAGCTAATTAGATCTTGT	23440
QY	395	-----ProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArg	410
DB	23441	TTTCCTTGACAGCCCTTTTCATAAATGGATGTTTCGACATATATACTTCCCGTCTGCGC	23500
QY	411	HisGlyLeuProLysAlaAla	417
DB	23501	AGCAAGATACCAAGAGGTGAG-TGAGATATATACCGATATGCAATTGTCGAGATTTGTTTC	23559
QY	418	-----AlaLeuLeuIleAla	422
DB	23560	TGTGATATAAATTTAACCTCCACACACTTGTTTTTCAGACACTCGCCATTAATCATTCGT	23619
QY	423	PheLeuValSerAlaLeuPheHisGlu-	431
DB	23620	TTCCCTAGTCTCTGCAGTCTTTCATGAGGTATACATACTTCTACATTTGCCCTGTCTCTAG	23679
QY	431	-----	431
DB	23680	ACGCATGAACACACGCTAGTGAAGAAATGCTTAATATTCAAAGCATGTTTTTACTTAAC	23739
QY	432	-----LeuCysIleAlaValProCysHisIlePhe	441
DB	23740	GATCTGTGTTACAAATTTCCCTTTTGACAGAGCTATGCATCGCAGTTCCTTGTGCTCTCTTC	23799
QY	442	LysLeuTrpAlaPheGlyGlyIleMetPheGln-	452

```
|||||
Db 23800 AAGCTATGGCTTTTCTTGGGATTATGTTTCAGGTTAAAAAATTACTAAACTGCTGCAGT 23859
Qy 453 -----ValPro 454
|||||
Db 23860 CGATTTTAAACTCTAATCTCATATTCGACCAACCAATTTGTTGAGTAGGTGCCT 23919
Qy 455 LeuValLeuIleThrAsnTyrLeuGlnAsnLysPhe--ArgAsnSerMet----- 470
||||| ||||||| ||||||| :||| |||
Db 23920 TTGGTCTTCATCACAAACTATCTACAGAAAGTTTGGCTCAACGGTATGCTCTCAAAAC 23979
Qy 470 ----- 470
|||||
Db 23980 CCGAGAAAATAGAACGAATAACTCTTCTTCATAGCCTAGCCATTTAAATCGCAATGCT 24039
Qy 471 -----ValGlyAsnM 474
|||||
Db 24040 GAAACTTAATAAAGTGATCTGTTTGAATGGGATCATATATTAGTGGGGAACA 24099
Qy 474 etIlePheTrpPheIlePheSerIleLeuGlyGlnPrometCysValLeuLeuTyrTyrH 494
||||| ||||||| ||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 24100 TGATCTTCTGTTTCATCTCTCGCATTTTCGGACAAACCGATGTGTGCTTCTTTATTACC 24159
Qy 494 IsAspLeuMetAsnArgLysGlyLysLeu 503
||||| ||||||| ||||||| ||| :|||
Db 24160 ACGACTGATGAACGAAAGATCGATG 24188
|||||

RESULT 14
AC003058 97495 bp DNA linear PLN 11-MAR-2002
LOCUS Arabidopsis thaliana chromosome 2 clone F27F23 map C1C06E08,
DEFINITION complete sequence.
ACCESSION AC003058
VERSION AC003058.3 GI:20197007
KEYWORDS HTG.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsia.
REFERENCE 1 (bases 1 to 97495)
AUTHORS Rounsley,S.D., Lin,X., Ketchum,K.A., Crosby,M.L., Brandon,R.C.,
Sykes,S.M., Kaul,S., Mason,T.M., Kerlavage,A.R., Adams,M.D.,
Somerville,C.R. and Venter,J.C.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 97495)
AUTHORS Lin,X.
JOURNAL Direct Submission
TITLE Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr., Rockville, MD 20850, USA
REFERENCE 3 (bases 1 to 97495)
AUTHORS Town,C.D. and Kaul,S.
JOURNAL Direct Submission
TITLE Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr., Rockville, MD 20850, USA, cdtown@igir.org
COMMENT On Apr 18, 2002 this sequence version replaced gi:6598388.
FEATURES
source
location/qualifiers
1. 97495
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/map="C1C06E08"
/clone="F27F23"
misc_feature
1. 9642
(note="overlap with BAC clone T20K24
(AC002392:95887..105528).")
complement(21..3689)
gene
/clone="F27F23"
/notes="F27F23.1"
mRNA
complement(join(<21..429,506..695,770..832,927..1053,
1129..1340,1428..1594,1710..1790,1882..1953,2038..2109,
2186..2324,2406..2884,2963..3509,3602..3689))
```

```
/gene="At2g19210"
complement(join(21..429,506..695,770..832,927..1053,
1129..1340,1428..1594,1710..1790,1882..1953,2038..2109,
2186..2324,2406..2884,2963..3509,3602..3689))
/gene="At2g19210"
/codon_start=1
/product="putative receptor-like protein kinase"
/protein_id="AAC16451.1"
/db_xref="GI:3135251"
/translation="MVHYNFSLIIFACFFAVFVLLVRAQDSGFVSIDGIPEDSSY
NDETDTIKYSDAAFAVESGIIHSIDPEFQSSLEKQFQNVRSFPEGNNRCDVKKPPQG
KGKYLIRTRFMYGNLDNKGAPDFDLYLGFNIWDSVTIDNATTIVTKIETLRSDH
HVCLVDKNGRTFPLSALEIRLLKSNITYETPYDSLILFKRMDLGGALGALVRYKDDYF
DRWIPLRFPKTYIFNASLTDNNNEQFQFARFVMTATSPEDLSODIIFSWEPKDP
TWKYFVYMHEAVVELSPNETREFKVLNKEINMSSFSRPLYLTDLTFLVQNPVSGPK
LEPRLOQTPRSTPLPIINATETVYVNEFLOSPTDQDDVDAMIRIKSKYVKSWLGD
CAPVKPDKDINGSYVDNESPRIISVNLSSGLTGEIDAFNSLTLHLIIDLVSNSUT
GKIPDFGNHNLTELNEGNKUSGAIPKVLKRSNKKLILRLDRIDGCVRAGPLDITKRY
SDEKTKNVIYIPLVASVGVGLVLAIALFLLYKKRRGGSGGVVRAGPLDITKRY
KYSEVVKVTNNFVRVGGQGVKGVHGVNDQVAVKILSESAQGYKEFRAEVELLL
RVHKNLTALIGYCHEGKMAIYEFMANGTLGDYLSGKSYVLSWELQISLDAAG
GLEVLHNGCKPPIVORDVKPANTILINEKLOAKIADGCLSRVALDGNNOPTAVAGTI
GYLDPVHLTKLSEKSDIYSFGVLLVYVSGQPVARSRTTAENIHTDRVDLMST
GDINGVDPKLGKRGFAGSAWKITEVAMACASSSSKNRPTMSHVVAELKESVSKRAG
GGSGASSVDPAMTNFDSGMPPQAR"
complement(5018..5047)
/rpt_family="At_rich"
complement(5060..6543)
/notes="F27F23.2; predicted by genscan"
complement(join(<5060..5590,5665..6171,6262..6543))
/gene="At2g19220"
complement(join(5060..5590,5665..6171,6262..6543))
/gene="At2g19220"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC16452.1"
/db_xref="GI:3135252"
/translation="MTSKAAWEPEHDEVFVLDLQVEQKMLGNQEMOHILEAFQEMGVY
FTDQLINHWDTMIKQKWCVRVQCKDKWDSLTNTFGATDQEWYLVNYPESAGY
RNPPLLEKLEIIFAGMNDLDEGTSKGMKQICEDREENVTVGVPRLSADIDTR
RHYKWPSSHAIVVDCFOBSLAKGIRPKRHLFTKESWKKILEKINRTGLGTYTHQ
LEHNFTRTRTSWAKHCTEIASPIKMDANTRRKFGATEDWDKLYLMINKRVRKRRHI
SDDRAEDGNVEPTSLRSDSEDAEVTPEMEKIPVNASVKKKEKFTFEECECLDA
IEEVERKGDVMEFLDKFYRLFMLQKSLRMAWLLKLSNP"
complement(5603..5623)
/rpt_family="At_rich"
complement(9141..12813)
/gene="At2g19230"
/notes="F27F23.3"
complement(join(<9141..9558,9652..9841,9925..9987,
10073..10199,10282..10478,10565..10722,10838..10918,
11004..11075,11166..11237,11314..11452,11535..12019,
12091..12637,12729..12813))
/gene="At2g19230"
complement(join(9141..9558,9652..9841,9925..9987,
10073..10199,10282..10478,10565..10722,10838..10918,
11004..11075,11166..11237,11314..11452,11535..12019,
12091..12637,12729..12813))
/gene="At2g19230"
/codon_start=1
/product="putative receptor-like protein kinase"
/protein_id="AAC16453.1"
/db_xref="GI:3135253"
/translation="MGNFNPFLVPSFASVYVLLVCAQDSGFVSIDGIPEDSSYY
DEKDTIKYISDAAFVESGIIHSIDSKFQKNLEKQFQNVRSFPEGNNRCDVKKPPQK
GKYLIRTRFMYGNLDNKGAPDFDLYLGFNIWDSVTLENSTIVTKIETLRSDK
HVCLVDKNGRTFPLSALEIRLLKSNITYETPYDSLILFKRMDLGGALGALVRYKDDYF
RWMPLMFNPFLINLTSLMIDPTSSNGFTPPSVVMYTRAWDLGATGLDLPARYKDDIDF
NKWFYIIFHAEEVKLPSNTRREFSVFLANKEQIDTTSVSPSYLYTDTLTVQNPVSGP
FLFELVRQGVKSRPPIPMNAETIYRINEFLDITDQNDVDAMIKITKYVKKNLWGLD
PCAPGTPMGGINCSTIANNPPRIISVNLSSGLTGQIDVPFITLPLQLKLDLSNNHL
```



Qy 157 -----CysCysLeuSe 160  
Db 91586 ATAATGATGAGTTAGTGATCAAGTTCCTCTTTGATTCTCTCTTCTGTCAGTATATC 91645  
Qy 160 rLeuValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIle 180  
Db 91646 CTTTCGATCTTCTTTGGTGGCTTTACGGTTGAGANAATGGTACTTCAGANAATACAT 91705  
Qy 180 eProGluProVal-ValValVal----- 187  
Db 91706 ATCAGAACCTGTGAGTAATATTACTATCCAGCCATTACTGTAATTTTATTGAAGACAA 91765  
Qy 188 -----L 188  
Db 91766 GTTTGTATCATGAAGAACTTACAAGTTCTGTGTTTGAANAATGCTCAAGTTGTCATCTTTC 91825  
Qy 188 euHisIleIleIleThrSerLeuPheTyrProValLeuValIleLeu----- 205  
Db 91826 TTCATATTATTATCACCATGACAGAGTTTGTATCCAGTTTACGTCACCCCTAAGGTGAT 91885  
Qy 205 ----- 205  
Db 91886 ACTGTTTCTGCTCAGTTTGTGATACTGTTTTTAAGTTTGTCTGACCCGGGA 91945  
Qy 206 -----ArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPheS 221  
Db 91946 TCTTGAANAATGGACAGGTGATGCTGCTTTTATCAGGTGTCATCTGATGCTCTCA 92005  
Qy 221 erCysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyrAspMetArgAlaL 241  
Db 92006 CTTGATGTGTGCTAAGTTGTTCTTATGCTCATATCTAGCTATGACATAGAATCCC 92065  
Qy 241 euThrLysLeuValGluLys-----GlyGluAlaLeuLeuAspThr 254  
Db 92066 TAGCAATGCAGCTGATGAAGTAAATACGAAAGAGCGTATGTAATT-AGTCACCTGC 92124  
Qy 255 -----LeuAsnMetAspTyrProT 261  
Db 92125 ACTGTGTACTGTTTTAACCAAAACACTGTTATGAACTTTAGGCCAATCCTGAAGCTCTCT 92184  
Qy 261 yrAsnValSerPheLysSerLeuAlaTyrPheLeuValAlaProThrLeuCysTyr----- 279  
Db 92185 ACTAGTTAGCTTGAAGAGCTTGGCATATTTTCATGGTCTGCCATCTGTTTATCAGG 92244  
Qy 279 ----- 279  
Db 92245 TAACTGCAAGTGCATCAACCATCTTATACTTGAAGAGTTTCTGTCTAAACCTCGGA 92304  
Qy 280 -----GlnProSerTyrProArgThrProTyrIleArgLysGlyTrpLeuP 295  
Db 92305 TCTTTTGTCTTTCCCGCCAGCAAGTTTATCCAGCTTCTGCATGTATACGGAAGGTTGGGTGG 92364  
Qy 295 heArgGlnLeuValLysLeuIleIlePheThrGlyValMetGlyPheIleIleAsp----- 313  
Db 92365 CTGTCATTTTGCARAACTGGTCATATTCACCGGATTCATGGGATTTATATAAGAACAG 92424  
Qy 313 ----- 313  
Db 92425 TACGTTTTCACATCTGTGCTTATTAGTTTCTCTTGGTGAAATCATCATCCCTGCGTTGT 92484  
Qy 314 -----GlnTyrIleAsnProIleValGln 321  
Db 92485 CACCACCTTGACTTCATGCTCTTTTGTACATTTTGGCAGTATATAAATCCTTATTGTGAGG 92544  
Qy 322 AsnSerGlnHisProLeuLysGlyAsnLeuLeuTyrAlaThrGluArgValLeuLysLeu 341  
Db 92545 AACTCAAGCATCTTTGAAGGCGATCTCTATATGCTATTGAAGAGTGTGAAGCTT 92604  
Qy 342 SerValProAsnLeuTyrValTrpLeuCysMetPheTyrCysPhePheHisLeuTrp----- 360  
Db 92605 TCAGTTCCAAATTTATATGCTGCTCTGTCATGTTCTACTGCTCTTCTCCACCTTTGCTAT 92664  
Qy 360 ----- 360

Db 92665 GCTGTGATCCCATCTCTTTCAAAATAATTTGCAAAATTCGAAAAACCGAAAAAGGCTAAAT 92724  
Qy 360 ----- 360  
Db 92725 CTCATACGAATTTGATATATTTTAGTTCTTTAGAGTCGGTGATGTAATTTTCACTACTGAA 92784  
Qy 361 -----LeuAsnIleLeuAlaGluLeuLeuArgPheGlyAspArg 373  
Db 92785 CGCAAACTCTCTTGTCCAAAGGTTAAACATATTGCAGAGCTTCTCTGCTCGGGGATCGT 92844  
Qy 374 GluPheTyrLysAspTrpTrpAsnAlaLysThrValGlu----- 386  
Db 92845 GAATTCACAAAGATGTGGGAATGCAGAAAGTGTGGG-AGATGTGAGCTATTTTACTCA 92903  
Qy 386 ----- 386  
Db 92904 AAAGAAAACATTATGATTTTAAATGTTGCTGTTGTTTTGGGTCACTTAACTAACCAAAATT 92963  
Qy 387 -----AspTyrTrpArgMetTrpAsnMet----- 394  
Db 92964 CATGATTTCACTGCTTCTCTTTATCAGTACTGAGAAATGTGGAATATGGTATGTTCTCT 93023  
Qy 394 ----- 394  
Db 93024 TCCTAAACATCACCTTCTTTTGTACAAAAATAGAAGAGAGACTAATTAAGATCTTGT 93083  
Qy 395 -----ProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArg 410  
Db 93084 TTTCTCTGACAGCTGTTCATAAATGGATGGTTCGACATATATATCTTCCCGCTGCTGGC 93143  
Qy 411 HisGlyLeuProLysAlaAla----- 417  
Db 93144 AGCAAGATACCAAGGTCAG-TGAGATATATACCGATATGCAATGTGCGAGATTTGTTTC 93202  
Qy 418 -----AlaLeuLeuIleAla 422  
Db 93203 TGTGATATAAATTTAACCCCTCCACACACTTGTGTTTTTCAGACACTCGCCATTTATCATGCT 93262  
Qy 423 PheLeuValSerAlaLeuPheHisGlu----- 431  
Db 93263 TTCTAGTCTCTGAGCTCTTCATGAGGTATACATACTTTCTACATTTGCCCTGTCTAG 93322  
Qy 431 ----- 431  
Db 93323 ACGCATGAACACACAGCTAGTGAAGAAATGCTAATATTCAAAGCATTTGTTTTACTTAAC 93382  
Qy 432 -----LeuCysIleAlaValProCysHisIlePhe 441  
Db 93383 GATCTTGTGTACAAATTTCTTTTGACAGCTATGTCATCGCAGTTCTCTTGTCTCTTTC 93442  
Qy 442 LysLeuTrpAlaPheGlyGlyIleMetPheGln----- 452  
Db 93443 AAGCTATGGGCTTTCTTGGGATTTATGTTTCAGGTTAAAAAATTTACTAAACTGCTGCAGT 93502  
Qy 453 -----ValPro 454  
Db 93503 CGATTTTACTTAAACTCTAATCTCATATCTCATATTCGACCAACCAATTTGTTTGAAGTGCCT 93562  
Qy 455 LeuValLeuIleThrAsnTyrLeuGlnAsnLysPhe-ArgAsnSerMet----- 470  
Db 93563 TTGCTCTTCATCAAACTATCTACAGAAAGGTTTGGCTCAACGGTATGCTCTCAAAAC 93622  
Qy 470 ----- 470  
Db 93623 CCGAGAAAATAGAAGAAATAACTCTTCTTTTCATAGCTAGCCATTTAAATCGCAATGCT 93682  
Qy 471 -----ValGlyAsnM 474  
Db 93683 GAAACTTAATAATAAAGGTGATCTGTTTGGAAATGGGATCATATTAATTAGTGGGGAACA 93742  
Qy 474 etIlePheTrpPheIlePheSerIleLeuGlyGlnProMetCysValLeuLeuTyrTrpH 494  
|||||

Db 93743 TGATCTCTGGTTCATCTTTCGATTTTCGACAAACCGATGCTGTGCTCTTATTATACC 93802  
Qy 494 isaspLeuMetAsnArgLysGlyLysLeu 503  
|||||  
Db 93803 ACGACGTGATGAACCGAAAGGATCGATG 93831  
RESULT 15  
AF078752 1650 bp mRNA linear ROD 11-NOV-1998  
DEFINITION Mus musculus diacylglycerol acyltransferase (Dgat) mRNA, complete cds.  
ACCESSION AF078752  
VERSION AF078752.1 GI:3859933  
KEYWORDS Mus musculus.  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 1650)  
AUTHORS Cases, S., Smith, S.J., Zheng, Y.-W., Myers, H.M., Lear, S.R., Sande, E., Novak, S., Collins, C., Welch, C.B., Lusis, A.J., Erickson, S.K. and Farese, R.V. Jr.  
TITLE Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (22), 13018-13023 (1998)  
MEDLINE 9789033  
PUBMED 9789033  
REFERENCE 2 (bases 1 to 1650)  
AUTHORS Farese, R.V. Jr.  
TITLES Direct Submission  
JOURNAL Submitted (16-JUL-1998) Gladstone Institute, 2550 23rd Street, San Francisco, CA 94110, USA  
FEATURES  
source Location/Qualifiers  
1..1650  
/organism="Mus musculus"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/chromosome="15"  
1..1650  
/gene="Dgat"  
96..1592  
/gene="Dgat"  
/note="Dgat"  
/codon\_start=1  
/product="diacylglycerol acyltransferase"  
/protein\_id="AAC72917.1"  
/db\_xref="GI:3859934"  
/translation="MGDRGGAGSSRRRTGSRSVGGGPKVEDEVRDAVSPDLG  
AGGAPAPAPAPATRDKGRTSVGDGVYDLRCHPLQDSLSSDSFGSNYKILNWCY  
VMLILSNARPLENLIKYLVDPIQVYSLFLKDPYSNPAPCVIIASNIFVVAFOIE  
KRLAVGALTEQMLLHVNLATIIICFPAVALIVESIPDNLTYRDLIYFIFAPTLCYEL  
SYRDVNLRCRRVRKAKAVSTGKRVSGAAQAQVSYPDNLTYRDLIYFIFAPTLCYEL  
NPPSPRIKRFLLRRVLEMLFFQTQVGLIQQMVPTIQNSMKPFKMDYSRIERL  
LKLAVPHLILWLIFFYWFHSCINAVAELOFGDFREYRDWMAESVYFQWNNIPV  
HKWCIIRHYKPLMRHSGSKWARTGVLTSAFFHELYLSDPLRMFLNAFTAMMAQVP  
LAWIVREFQNGYCNAAVWVLIIGQPVAVLMVYHDYVLYNDAFVGV"  
BASE COUNT 333 a 427 c 471 g 419 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,46e-62 Length: 1650  
Score: 807.50 Matches: 190  
Percent Similarity: 55.36% Conservative: 89  
Best Local Similarity: 37.70% Mismatches: 174  
Query Match: 30.13% Indels: 51  
DB: 10 Gaps: 16  
US-09-856-018B-16 (1-504) x AF078752 (1-1650)  
Qy 17 SerSerLeuArgArgPro----SerAlaThrSerThrAlaGlyLeuPheAsnSerPro 35  
|||||  
Db 120 AGCTCTCGCGCTCGGAGACCGCGCTCGGGGTTTCGCTCCAGGT-----164

Qy 36 GluThrThrThrAspSerSerGlyAspAspLeuAlaLysAspSerGlySerAspSer 55  
|||||  
Db 165 -----GGTAGTGGGCCCAAGGTAGAGAGGACGAGGTGGGACGCGGCT 209  
Qy 56 IleAsnSerAsp-----AspAlaAlaValAsnSerGlnGlnAsnGlu 70  
|||||  
Db 210 GTGAGCCCGGACTTGGGCGCGGGGTGACGCGCGGCTCCGGCTCCAGGCCAT 269  
Qy 71 LysGlnAspThrAspPheSerValLeuLysPheAlaTyrArgProSerVal----- 87  
|||||  
Db 270 ACCCGGACAAAGAC-----GGCGGACCAAGCGCTGGGCGACGGC 308  
Qy 88 -----ProAlaHisArgLysValLysGluSerProLeuSerSerAspPheThr 103  
|||||  
Db 309 TACTGGGATCTGAGGTGCCATCGT-----CTGCAAGATCTTCTTTCAGCTCAGACAGTGGT 365  
Qy 104 PheArgGlnSerHisAlaGlyLeuPheAsnLeuLysValLeuValAlaValAsn 123  
|||||  
Db 366 TTC---AGCAATATCTGCTGATCTCGAATTTGGTGTGTGATGCTGATCTCGAGTAAT 422  
Qy 124 SerArgLeuIleGluAsnLeuMetLysTyrGlyTrpLeuLysSerGlyPheTyr 143  
|||||  
Db 423 GCAAGGTATTTTAGAGAACCTTATCAAGTATGATGCTCTGGTG---GATCCTATCCAG 479  
Qy 144 PheSerSerLysSerLeuArgasp-----TrpProLeuPheMetCysCysLeuSer 160  
|||||  
Db 480 GTGGTGTCTCTGTTTGAAGGACCCCTACAGCTGCCCTGCCCATGCGGTGATTATGCA 539  
Qy 161 LeuValValPheProPheAlaAlaPheIleValGluLysLeuAlaGlnArgLysCysIle 180  
|||||  
Db 540 TCCAATATTTTGTGTGCTGCTATTCAGATTTGAGAACGCGCTGGCAGTGGTGGCCCTG 599  
Qy 181 ProGluProValValValLeuHisIleIleIleThrSerThrSerLeuPheTyrPro 200  
|||||  
Db 600 ACAGAGCAGATGGGCTGCTGTACATGTGGTTAACTGGCCACATCATTTGCTCCCA 659  
Qy 201 ValLeuValIleLeuArgCysAspSerAlaPheValSerGlyValThrLeuMetLeuPhe 220  
|||||  
Db 660 GCAGCTGTGGCCTTACTGTTGAGTCTATCACTCCAGTGGGTTCCTGTTGCTCGGCA 719  
Qy 221 Ser---CysValValTrpLeuLysLeuValSerTyrAlaHisThrAsnTyr----- 236  
|||||  
Db 720 TCATCTCCATCATGTCTCTCAAGCTTTATTCCTACCGGAGTCAACCTGGTGGTCCGC 779  
Qy 237 AspMetArgAlaLeuThrLysLeuValGluLysGlyGluAlaLeuAspThrLeu--- 255  
|||||  
Db 780 CAGCGAAGGCTCAAGCCCAAGCTGCTCTACAGGAAGAAGGTCACTAGTGGGCTGCTGCC 839  
Qy 256 -----AsnMetAspTyrProTyrAsnValSerPheLysSerLeuAlaTyrPheLeuVal 273  
|||||  
Db 840 CAGCAAGCTGTGAGCTATCCAGAACACCTGACCTACCGAGATCTCTATTACTTCACTTT 899  
Qy 274 AlaProThrLeuCystyrGlnProSerTyrProArgThrProTyrIleArgLysGlyTyr 293  
|||||  
Db 900 GCTCCTACTTTGTTTATGAACCTCACTTCTCGGTCCTCCCGCAATACAGAACGCTTT 959  
Qy 294 LeuPheArgGlnLeuValLysLeuIlePheThrGlyValMetGlyPheIleLeuAsp 313  
|||||  
Db 960 CTGCTACGACGAGTCTCTGAGATGCTCTTTTACCCAGCTTCAAGTGGGCTGATCCAA 1019  
Qy 314 GlnTyrIleAsnProIleValGlnAsnSerGlnHisProLeuLysGlyAsnLeuTyr 333  
|||||  
Db 1020 CAGTGGATGTCCTACTATCCAGAACCTCCATGAAGCCCTTCAAG---GATATGGACTAT 1076  
Qy 334 Ala-----ThrGluArgValLeuLysLeuSerValProAsnLeuTyrValTrpLeuCys 351  
|||||  
Db 1077 TCACGATCATTTGAGGCTCTTTAAAGCTGGCGGTCCCAACCATCTGATCTGGCTATC 1136  
Qy 352 MetPheTyrCysPhePheHisLeuTrpLeuAsnIleLeuAlaGluLeuArgPheGly 371  
|||||  
Db 1137 TTCTTCTATTTGTTTTCACCTCTCTCAATGCTGTGGCAGAGCTTCTGCAAGTTGGA 1196  
Qy 372 AspArgGluPheTyrLysAspTrpTrpAsnAlaLysThrValGluAspTyrTrpArgMet 391

```
|||||
Db 1197 GACCGCGAGTTCTACAGAGATTGGTGGAACTGCTGAGTCTGCACCTACTTTTGGCAGAAC 1256
Qy 392 TrpAsnMetProValHisLysTrpMetIleArgHisLeuTyrPheProCysLeuArgHis 411
|||||
Db 1257 TGGAAATATCCCGCGCACAACTGGTGCATCAGACACTTCTACAAGCCCTATGCTCAGACAT 1316
Qy 412 GlyLeuProLysAlaAlaAlaLeuLeuIleAlaPheLeuValSerAlaLeuPheHisGlu 431
|||||
Db 1317 GGCAGCAGCAAAATGGGTGGCCAGGACAGGAGTATTTTGTACCTCAGCCTTCTTCATGAG 1376
Qy 432 LeuCysIleAlaValProCysHisIlePheLysLeuTrpAlaPheGlyGlyIleMetPhe 451
|||||
Db 1377 TACCTAGTGAGCGTTCCTCCCTCGCGATGTTCCGCTCTGGGCATTTCACAGCCATGATGGCT 1436
Qy 452 GlnValProLeuValLeuIleThrAsnTyrLeuGlnAsnLysPheArgAsnSerMetVal 471
|||||
Db 1437 CAGGTCCCACTGGCC-----TGGATTGTGGCCCGATTCTTCCAAGGGAACATAT 1484
Qy 472 GlyAsnMetIlePheTrpPheIlePheSerIleLeuGlyGlnPrometCysValLeuLeu 491
|||||
Db 1485 GGCAATGCAGCTGTGTGG--GTGACACTCATCATTTGGGCAACCGGTGGTGTGTGTCATG 1541
Qy 492 TyrTyrHisasp 495
|||||
Db 1542 TATGTCCACGAC 1553
```

Search completed: February 21, 2003, 20:49:52  
Job time : 3493 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 15:29:03 ; Search time 430 Seconds  
(without alignments)  
10170.653 Million cell updates/sec

Title: US-09-856-018b-15  
Perfect score: 1942  
Sequence: 1 tagaaacacactcggtctt.....aaaaaaaaaaaaaaaaaaaa 1942

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				N_Geneseq_101002.*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*						
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*						
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*						
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*						
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*						
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*						
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*						
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*						
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*						
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*						
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*						
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*						
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*						
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*						
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*						
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*						
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*						
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*						
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*						
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*						
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*						
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*						
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*						
24:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1942	100.0	1942	21 AAA48939	Soybean diacylglyc
2	719.8	37.1	1904	21 AAA51482	A. thaliana diacyl
3	719.8	37.1	1942	21 AAA88835	Arabidopsis acyl C
4	719.8	37.1	1942	21 AAZ45371	Acyl-CoA:cholester
5	719.8	37.1	1942	22 AAS01106	Arabidopsis thalia
6	716.6	36.9	1888	21 AAA48932	Arabidopsis diacyl
7	683	35.2	1985	21 AAA51484	A. thaliana AS11 d
8	672.8	34.6	1975	21 AAA48942	Wheat diacylglycer
9	630.2	32.5	1587	21 AAA48938	Rice diacylglycero

10	478.8	24.7	1281	21	AAA48933	Corn diacylglycerol
11	413.4	21.3	901	21	AAA48936	Corn diacylglycerol
12	404.8	20.8	1559	21	AAA48935	Corn diacylglycerol
13	259.6	13.4	470	21	AAA48940	Soybean diacylglycerol
14	253.2	13.0	380	24	ABQ85322	Arabidopsis thaliana
15	235.2	12.1	978	21	AAA48934	Corn diacylglycerol
16	229.8	11.8	629	21	AAA51486	EST with homology
17	229.8	11.8	629	21	AAZ49453	A. thaliana diacylglycerol
18	199.4	10.3	234	21	AAA88839	Soybean acyl CoA:cholesterol
19	199.4	10.3	234	21	AAZ45372	Acyl-CoA:cholesterol
20	199.4	10.3	234	22	AA501311	Soybean sterol acyl CoA:cholesterol
21	194.6	10.0	275	21	AAA88840	Soybean acyl CoA:cholesterol
22	194.6	10.0	275	21	AAZ45374	Acyl-CoA:cholesterol
23	194.6	10.0	275	22	AA501313	Soybean sterol acyl CoA:cholesterol
24	187.2	9.6	267	21	AAA88841	Soybean acyl CoA:cholesterol
25	187.2	9.6	267	21	AAZ45373	Acyl-CoA:cholesterol
26	187.2	9.6	267	22	AA501312	Soybean sterol acyl CoA:cholesterol
27	170.2	8.8	1650	21	AAZ49452	Mouse Diacylglycerol
28	162.6	8.4	1521	19	AAV01533	Human acyl CoA:cholesterol
29	162.6	8.4	1976	21	AAZ76169	Human ACAT Related
30	159.8	8.2	1766	21	AAA88846	Rat acyl CoA:cholesterol
31	159.8	8.2	1766	21	AAZ45385	Acyl-CoA:cholesterol
32	159.8	8.2	1766	22	AA501105	Rat sterol acyltransferase
33	159.4	8.2	1895	21	AAA88842	Human acyl CoA:cholesterol
34	159.4	8.2	1895	21	AAZ45383	DNA encoding a protein
35	154	7.9	993	22	ABA06413	Human cDNA SEQ ID
36	149.6	7.7	470	24	ABU93444	Arabidopsis thaliana
37	140.6	7.2	983	19	AAU01539	Human acyl CoA:cholesterol
38	130.4	6.7	254	21	AAA48937	Rice diacylglycerol
39	125.2	6.4	452	24	ABL81686	Human ovarian cancer
40	124.4	6.4	253	21	AAA88849	Maize acyl CoA:cholesterol
41	124.4	6.4	253	21	AAZ45376	Acyl-CoA:cholesterol
42	124.4	6.4	253	22	AA501315	Maize sterol acyltransferase
43	122.4	6.3	3996	23	ABL28621	Drosophila melanogaster
44	121.6	6.3	646	21	AAA48941	Wheat diacylglycerol
45	121.2	6.2	1122	20	AAU30335	DNA encoding a human protein

ALIGNMENTS

RESULT 1			
AAA48939			
ID	AAA48939	standard; cDNA; 1942 BP.	
XX	AC	AAA48939;	
XX	DT	06-DEC-2000 (first entry)	
XX	DE	Soybean diacylglycerol acyltransferase cDNA #1.	
KW	Diacylglycerol acyltransferase; corn; rice; soybean; wheat;		
KW	triacylglycerol; herbicide; EC2.3.1.20; ss.		
XX	OS	Glycine max.	
XX	XX	Key	Location/Qualifiers
FT	CDS	29..1543	
FT	FT	/tag=	a
FT	FT	/product=	Diacylglycerol_acyltransferase
XX	PN	WO200032756-A2.	
XX	PD	08-JUN-2000.	
XX	PF	01-DEC-1999;	99WO-US28354.
XX	PR	02-DEC-1998;	98US-0110602.
XX	PR	31-MAR-1999;	99US-0127111.
XX	PA	(DUPO.) DU PONT DE NEMOURS & CO E. I.	
XX	XX	Cahoon EB, Kinney AJ, Cahoon RE;	

XX WPI; 2000-412308/35.  
DR P-PSDB; AAY94519.  
XX  
PT Polynucleotides encoding diacylglycerol acetyltransferase, useful for  
PT synthesis of triacylglycerols and increasing the level of oils in plant  
XX seeds -  
XX  
PS Claim 4; Page 51; 62pp; English.  
XX  
CC In the present invention, cDNA libraries from Arabidopsis, corn, rice,  
CC soybean and wheat were screened for sequences with homology to a  
CC putative acyl CoA cholesterol acyltransferase related gene from  
CC Arabidopsis thaliana and diacylglycerol acyltransferases from Homo  
CC sapiens and Mus musculus. The cDNA clones identified from this process  
CC were used to form complete diacylglycerol acyltransferase cDNA  
CC sequences. The present sequence is soybean diacylglycerol  
CC acyltransferase cDNA from clone srl.pk0098.a8. Diacylglycerol  
CC acyltransferases are involved in the synthesis of triacylglycerols.  
CC Alteration of the expression of the diacylglycerol acyltransferase  
CC DNA can be useful for increasing the level of oils in plant seeds.  
CC Inhibitors of diacylglycerol acyltransferase may be useful as  
CC herbicides.  
XX  
SQ Sequence 1942 BP; 517 A; 425 C; 377 G; 623 T; 0 other;

Query Match 100.0%; Score 1942; DB 21; Length 1942;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1942; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAGAACAACACGCTGGTCTTCTTCCAAATGGCGATTTCCGATGAGCCTGAAAGTGATGC 60  
Db 1 TAGAACAACACGCTGGTCTTCTTCCAAATGGCGATTTCCGATGAGCCTGAAAGTGATGC 60

Qy 61 CACTGCTCTCAACCACTCTCCCTGCGCGCGGCTCCCTCCGCGCCACTCCACGCGGCGCT 120  
Db 61 CACTGCTCTCAACCACTCTCCCTGCGCGCGGCTCCCTCCGCGCCACTCCACGCGGCGCT 120

Qy 121 CTTCAAATTCGGCTGAGACAACACCGACAGTTCGGGTGATGACTTGGCCAAAGGATTCG 180  
Db 121 CTTCAAATTCGGCTGAGACAACACCGACAGTTCGGGTGATGACTTGGCCAAAGGATTCG 180

Qy 181 TTCGACGACTCCATCAACAGCGAGGAGCGCGCGCTCAATTCGCAACAGCAAAACGAAAA 240  
Db 181 TTCGACGACTCCATCAACAGCGAGGAGCGCGCGCTCAATTCGCAACAGCAAAACGAAAA 240

Qy 241 ACAAGACACTGATTTCCGCTCCAAATTCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 300  
Db 241 ACAAGACACTGATTTCCGCTCCAAATTCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 300

Qy 301 CAAAGTGAAGAAAGTCGGCTCAGCTCCGACACTATTTCCGTCAGAGTCACGCGGCGCT 360  
Db 301 CAAAGTGAAGAAAGTCGGCTCAGCTCCGACACTATTTCCGTCAGAGTCACGCGGCGCT 360

Qy 361 CTTCAACCTTTGTATAGTAGTCCCTTGTGCTGTAATAGCGCAGCTCATCTGAGAATTT 420  
Db 361 CTTCAACCTTTGTATAGTAGTCCCTTGTGCTGTAATAGCGCAGCTCATCTGAGAATTT 420

Qy 421 AATCAAGTATGTTGGTTGATCAAAATCGGCTTTTGGTTAGTTCAAAAGTCATTGAGAGA 480  
Db 421 AATCAAGTATGTTGGTTGATCAAAATCGGCTTTTGGTTAGTTCAAAAGTCATTGAGAGA 480

Qy 481 CTGCCCCCTTTTCATGTTGTTCTTCTTGTGTTGTTTCTTTCGCTGCTCTTATAGT 540  
Db 481 CTGCCCCCTTTTCATGTTGTTCTTCTTGTGTTGTTTCTTTCGCTGCTCTTATAGT 540

Qy 541 GGAGAAGTTGGCAACAGGAAGTGTATACCGAACCAAGTGTGTTGTTGTTGTTGTTGTT 600  
Db 541 GGAGAAGTTGGCAACAGGAAGTGTATACCGAACCAAGTGTGTTGTTGTTGTTGTTGTT 600

Qy 601 CATTAACCTCAACTTTCGCTTTTCTATCAGATTTTATCTCAGGCTGATTCGCTTTT 660  
Db 601 CATTAACCTCAACTTTCGCTTTTCTATCAGATTTTATCTCAGGCTGATTCGCTTTT 660

Qy 661 TGATACAGTGTACAGTTAAATGCTGTTTCTGTTGTTGTTTAAATTTGGTGTCTTTA 720  
Db 661 TGATACAGTGTACAGTTAAATGCTGTTTCTGTTGTTGTTTAAATTTGGTGTCTTTA 720

Qy 721 TGCACATACAAACTATGATATGAGAGCACTTACCAAAATAGTTGAAAAGGAGGAGCACT 780  
Db 721 TGCACATACAAACTATGATATGAGAGCACTTACCAAAATAGTTGAAAAGGAGGAGCACT 780

Qy 781 GCTCGATACCTGAACATGGAGTATCTTACAACTAGCTTCAAGAGCTTGGCATATTT 840  
Db 781 GCTCGATACCTGAACATGGAGTATCTTACAACTAGCTTCAAGAGCTTGGCATATTT 840

Qy 841 CTTGTTGCCCTTACATATGATTTACCAAGCTATCTCGCACACCTTATATTCGAAA 900  
Db 841 CTTGTTGCCCTTACATATGATTTACCAAGCTATCTCGCACACCTTATATTCGAAA 900

Qy 901 GGGTGGTTGTTTCCGCAACTTGTCAAGCTGATATTTACAGAGTATATGGGATTTAT 960  
Db 901 GGGTGGTTGTTTCCGCAACTTGTCAAGCTGATATTTACAGAGTATATGGGATTTAT 960

Qy 961 AATAGACCAATATATTAATCCATAGTACAAAATTCACAGCATCTCTCAAGGGAACCT 1020  
Db 961 AATAGACCAATATATTAATCCATAGTACAAAATTCACAGCATCTCTCTCAAGGGAACCT 1020

Qy 1021 TCTTTAGCCCAACGAGAGTCTGAAGCTTCTGTTCCAAATTTATATGTTGGCTCTG 1080  
Db 1021 TCTTTAGCCCAACGAGAGTCTGAAGCTTCTGTTCCAAATTTATATGTTGGCTCTG 1080

Qy 1081 CATGTTCTATGCTTTTCCACCTTGGTTAAATATCTCGCAGAGCTTCTTCGATTTGG 1140  
Db 1081 CATGTTCTATGCTTTTCCACCTTGGTTAAATATCTCGCAGAGCTTCTTCGATTTGG 1140

Qy 1141 TGATCGTGAATTCACAGAGATTTGGTGAATGCCAAAATCTGCAAGATTTATTCGAGGAT 1200  
Db 1141 TGATCGTGAATTCACAGAGATTTGGTGAATGCCAAAATCTGCAAGATTTATTCGAGGAT 1200

Qy 1201 GTGGAATGATGCTGTTCACAAATGATGATCGGCACCTATATTTCCATGTTTAAAGCA 1260  
Db 1201 GTGGAATGATGCTGTTCACAAATGATGATCGGCACCTATATTTCCATGTTTAAAGCA 1260

Qy 1261 CCGTCTACAAAGGCTGCTGCTTTTAAATGCTTCTGCTTTCTGCTTTATTCCTATGA 1320  
Db 1261 CCGTCTACAAAGGCTGCTGCTTTTAAATGCTTCTGCTTTCTGCTTTATTCCTATGA 1320

Qy 1321 GCTGTGATTCGCTGCTTCCGACATATTCAGTGTGGGCTTTCGGTGAATATGTT 1380  
Db 1321 GCTGTGATTCGCTGCTTCCGACATATTCAGTGTGGGCTTTCGGTGAATATGTT 1380

Qy 1381 TCAGGTTCCCTTTGGTCTTGTGATCACTAAATTCGAAAAATAATTCAGAAACTCAATGGT 1440  
Db 1381 TCAGGTTCCCTTTGGTCTTGTGATCACTAAATTCGAAAAATAATTCAGAAACTCAATGGT 1440

Qy 1441 TGGAAATATGATTTTGGTTCATATTCAGTATCCTTGGTCAACCTATGTTGTGACTGCT 1500  
Db 1441 TGGAAATATGATTTTGGTTCATATTCAGTATCCTTGGTCAACCTATGTTGTGACTGCT 1500

Qy 1501 ATACTACCATGACTTGATAGTAAAGGCAAACTTGACTGAAGCTACGCCCATATACAT 1560  
Db 1501 ATACTACCATGACTTGATAGTAAAGGCAAACTTGACTGAAGCTACGCCCATATACAT 1560

Qy 1561 TTTAAAGTGTCACATGGATGAGCTTTTTCAGTTTTTCAGATTTGAAAATTTGATGTTGATG 1620  
Db 1561 TTTAAAGTGTCACATGGATGAGCTTTTTCAGTTTTTCAGATTTGAAAATTTGATGTTGATG 1620

Qy 1621 TTGTTCAATATTTGTTTTCAGAAATGCTTTCATCTACCATGGCAATTTGGTGTCTGAAG 1680  
Db 1621 TTGTTCAATATTTGTTTTCAGAAATGCTTTCATCTACCATGGCAATTTGGTGTCTGAAG 1680

Qy 1681 GAATTCACCGGATATGCCAGTTCACGAGGCTAATTCATTTATCTGATCTATGTTACTTAC 1740  
Db 1681 GAATTCACCGGATATGCCAGTTCACGAGGCTAATTCATTTATCTGATCTATGTTACTTAC 1740







QY 1201 GTGGAATATGCTGTTCAAAATGGATGATCCGCCACCTATATTTTCCATGTTTAAAGGCA 1260  
DB 1460 GTGGAATATGCTGTTCAAAATGGATGATCCGCCACCTATATTTTCCATGTTTAAAGGCA 1519  
QY 1261 CGGTCTACCAAAAGGCTGCTCTTTTAATTTGCTTCCCTGGTTCCTGCTTTATTCATGCA 1320  
DB 1520 CAAGATACCAAGACACTGCCATATCATGTTTCTTCTAGTCTCTGAGTCTTTTCATGA 1579  
QY 1321 GGTGTCATGCTGTTTCCATGACATATTCAGTTTGTGGGCTTTCCGGTGGAAATATGTT 1380  
DB 1580 GCTATGATCGGAGTTCCTTCTCTTCAAGCTATGGGCTTTTCTTGGGATATGTT 1639  
QY 1381 TCAGGTTCTTGGCTTGTATGATCACTAATTTATGCGAAATAAATTCAGAACTCATGTT 1440  
DB 1640 TCAGGTTCTTGGCTTGTATGATCACTAATTTATGCGAAATAAATTCAGAACTCATGTT 1696  
QY 1441 TGAATATGATTTTGTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1500  
DB 1697 GGGACATGATCTTCTGTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1756  
QY 1501 ATACTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1551  
DB 1757 TTATTACCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1807

RESULT 4

AZ45371  
ID AA245371 standard; DNA; 1942 BP.  
XX  
AC AA245371;  
XX  
DT 27-MAR-2000 (first entry)  
XX  
DE Acyl-CoA:cholesterol acyltransferase (ACAT)-like protein DNA.  
KW Acyl-CoA:cholesterol acyltransferase; ACAT; ACAT-like protein; sterol;  
KW ester; triacylglycerol; fatty acyl-CoA; lipid composition; plant cell;  
KW diacylglycerol acyltransferase; DAGAT; triglyceride; cancer; diabetes;  
KW cardiopulmonary disease; heart failure; atherosclerosis; adipocytosis;  
KW leukaemia; skin carcinoma; fibroblastoma; metabolic disorder; obesity;  
KW abnormal lipid metabolism; abnormal fat absorption;  
KW lipoprotein secretion; adipogenesis; ss.  
OS Arabidopsis thaliana.  
XX  
FH Key Location/Qualifiers  
FT CDS 237..1799  
FT /\*tag= a  
FT /\*product= "Acyl-CoA:cholesterol acyltransferase (ACAT)-  
FT like protein"

XX  
PN W09963096-A2.  
XX  
XX 09-DEC-1999.  
XX  
XX 04-JUN-1999; 99WO-US12541.  
XX  
XX 05-JUN-1998; 98US-0088143.  
PR 12-NOV-1998; 98US-0108389.  
XX  
XX (CALJ ) CALGENE LLC.  
XX  
XX Lassner MW, Ruzinsky DM;  
PI  
XX WPI; 2000-105701/09.  
DR P-PSDB; AAY54143.  
XX  
XX Novel polynucleotides used for modifying plant oil composition and for  
PT developing products for treating e.g. cancer, diabetes, cardiopulmonary  
PT disease or metabolic disorders  
XX  
XX Claim 4; Fig 1; 89pp; English.  
XX

CC The present sequence encodes an acyl-CoA:cholesterol acyltransferase  
CC (ACAT) related protein. The ACAT-like protein is active in the formation  
CC of a sterol, ester and/or triacylglycerol from a fatty acyl-CoA and  
CC sterol and/or diacylglycerol substrate. The DNA can be used for  
CC modifying the lipid composition of plant cells. The ACAT-like protein  
CC has diacylglycerol acyltransferase (DAGAT) activity, and so the  
CC synthesis of triglycerides can be suppressed or increased using the  
CC DNA. The protein can be used to produce plant oils with a modified  
CC triglyceride content. The products can also be used to identify  
CC antagonists and agonists of DAGAT activity. Such agonists and  
CC antagonists are particularly useful in treating or ameliorating  
CC diseases associated with DAGAT activity, including diseases associated  
CC with altered cellular diacylglycerol concentration or PKC activity,  
CC including cancer, diabetes, cardiopulmonary diseases e.g. heart failure,  
CC atherosclerosis, adipocytosis, leukaemia, skin carcinoma, fibroblastoma,  
CC metabolic disorders, obesity, diseases associated with abnormal lipid  
CC metabolism, and diseases associated with abnormal fat absorption,  
CC lipoprotein secretion and adipogenesis.  
XX  
SQ Sequence 1942 BP; 461 A; 421 C; 425 G; 635 T; 0 other;  
  
Query Match 37.1%; Score 719.8; DB 21; Length 1942;  
Best Local Similarity 73.2%; Pred. No. 3.5e-141;  
Matches 960; Conservative 0; Mismatches 327; Indels 24; Gaps 2;  
  
QY 241 ACAGACACTGATTTCTCGTCTCTCAAAATCGGCTACCGTCTTCTCGTCCCGCTCACCG 300  
DB 521 AGAAGGAAGGAACCGCGATGCTACGTTTACGTATCGACCGCTCGGTTCCAGCTCATCG 580  
QY 301 CAAAGTGAAGGAAGTCCGCTCAGCTCCGACACTATTTTCCGTCAGAGTCACGGGCGCT 360  
DB 581 GAGGGCAGAGAGAGTCCACCTAGCTCCGACGCAATCTTCAACAGAGCGCATGCCGGATT 640  
QY 361 CTTCAACCTTTGTATAGTAGTCTTCTGTTGTAATAGCCGACTCATCATGAGAAATTT 420  
DB 641 ATTCACCTCTGTAGTAGTCTTCTTATGCTGTAAACAGTAGACTCATCATCAAAATCT 700  
QY 421 ATGAAGTAGTGTGTTGATCAAAATCTGCTTTTGGTTTAGTTCAAGTCATTTAGAGAA 480  
DB 701 TATGAAGTAGTGTGTTGATCAAGAGGATTTCTGTTTAGTTCAAGTCATTTAGAGAA 760  
QY 481 CTGGCCCTTTTCATGTTGTTCTTCTCTGTTGTTATTTCTTCTGCTGCTGCTTTTATAGT 540  
DB 761 TTGGCCCTTTTCATGTTGTTATATCCCTTCGATCTTTCTTCTGCTGCTTACGGT 820  
QY 541 GGAGAAAGTTGGCAACAGGAAGTGTATACCGAACAGTGTGTTGTTGTTGTTGTTGTTAAT 600  
DB 821 TGAGAAATTTGGTACTTCAGAAATACATATACAGAACCTGTGTCATCTTCTTCATATAT 880  
QY 601 CATTACCTCAACTTCCTTTCTATCCAGTTTATGTTTATTTATTCAGGTGATGTTCTGTTT 660  
DB 881 TATCACCATGACAGAGGTTTGTATCCAGTTTACGTCACCCCTAAGGTGATGTTCTGTTT 940  
QY 661 TGTATCAGGTGTCACCTTAATGCTGTTTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720  
DB 941 TTTATCAGGTGTCACCTTTGATGCTCTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1000  
QY 721 TGCATACAAACTATGATGAGACACTTACCAAAATTTAGTTGAAAGGAGAGAACACT 780  
DB 1001 TGCTCATAGCTATGACATAGATCCCTAGCCATGACGCTGATAGG- - - - - 1050  
QY 781 GTCGATACCTGTAACATGAGACTATCCTTACAACGTAAGCTTCAAGAGCTTGGCATATTT 840  
DB 1051 - - - - - CCAATCCTCAAGTCTCCTACTACGTTAGCTTCAAGAGCTTGGCATATTT 1099  
QY 841 CTTGTTGCTCCCTACATTTATGTTACCGCAAGCTATCCTCGCACACCTTATATTCGAA 900  
DB 1100 CATGTCGCTCCACATTTGTTATCAGCAAGTTTATCCACGTTCTGATGTTACGGAA 1159  
QY 901 GGGTTGTTGTTTCGCAACTTGTCAAGCTGATAATTTTACAGAGGTTATGGATTTAT 960  
DB 1160 GGGTTGTTGTTCTGCAATTTCAAAACCTGTCATATTCACCGGATTCATGGGATTTAT 1219



[illegible]

11

```

FT
FT
/*tag= a

```

Db 835 TTTATCAGGTGTCACCTTTGATGCTCCTCCTACCTTGCAATGTTGGTGAAGTTGGTTTCCTTA 894  
Qy 721 TGCATACAAACTATGATATGAGACACTTACCAAAATAGTTGAAAGGGAGAACCACT 780  
Db 895 TGCTCATACTAGCTATGATGAAGATCCTAGCCCAATCAGCTGATAGG----- 944  
Qy 781 GCTGATACTCTGAACATGGACTATCCCTTACACAAAGCTTCAAGAGCTGGCAATATT 840  
Db 945 -----CCAATCTGAAGTCTCTCTACTACCTTACCTTGAAGAGCTGGCATATT 993  
Qy 841 CCTGTTCCCTACATATGTTACCAAGCTATCTGCAAGCTATCTGCAACCTTATATTCGAAA 900  
Db 994 CATGTCCTCCACATATGTTATGATGCAAGCTATCCAGCTTCTGCAATGATATACGAAA 1053  
Qy 901 GGGTGGTGGTTTCCCAACTTGTCAAGCTATATATTTACAGAGTATATGGGATTTAT 960  
Db 1054 GGGTGGTGGTCTGTCATTTGCAAACTGTCATATTCACCGATTCATGGGATTTAT 1113  
Qy 961 AATGACCAATATATTAATCCATAGTACAAATTCACAGCATCTCTCAAGGAAACCT 1020  
Db 1114 AATGACCAATATATAAATCTATGTCAGGAACCTCAAAAGCATCTTTGAAAGGCGATCT 1173  
Qy 1021 TCTTTACGCCACGAGAGAGTCTCAAGCTTCTGTTCCAAATTTATATGTTGGCTCTG 1080  
Db 1174 TCTATGCTATTTGAAGAGTGTGAAGCTTTCAGTTCCAAATTTATATGTTGGCTCTG 1233  
Qy 1081 CATGTTCTATGCTTTTCCACCTTGGTTAAATATCTGCGAGAGCTTCTCGAATTTGG 1140  
Db 1234 CATGTTCTACTGCTTCTTCCACCTTGGTTAAACATATTTGGCAGAGCTTCTGCTCGG 1293  
Qy 1141 TGATCGTAATCTACAGGATTTGGTGAATGCGCAAACTGTCGAAGATTTATGGAGAT 1200  
Db 1294 GGATCGTGAATCTACAAAGATTTGGTGAATGCGCAAACTGTCGAAGATTTATGGAGAT 1353  
Qy 1201 GTGAATATGCTGTTCACAAATGATCGCGCACTATATTTCCATCTTTAAGGCA 1260  
Db 1354 GTGNAATGCTGTTCATAATGATGGTTCGACATATATCTCCCGTCTGCGCAG 1413  
Qy 1261 CGGTCTACCAAGGCTGCTCTTTAAATGCTTCTGCTTTCTGCTTTATTTCCATGA 1320  
Db 1414 CAAGATACCAAGACACTCGCCATATATCTGCTTTCTCTAGTCTGCGAGCTTTCATGA 1473  
Qy 1321 GCTGTCATGCTGCTTCCGACATATTCAGTTGTTGGCTTTCGGTGAATATGTT 1380  
Db 1474 GCTATGATCGAGTCTCTGCTGCTCTTCAAGCTATGCGCTTTCTTGGGATATGTT 1533  
Qy 1381 TCAGTTCCTTTGCTTGTGATCACTAATATCTGCAAAATAAATTCAGAACTCAATGTT 1440  
Db 1534 TCAGTTCCTTTGCTTGTGATCACTAATATCTGCAAAATAAATTCAGAACTCAATGTT 1590  
Qy 1441 TGAATATGATTTTTTGGTTTCATATTCAGATCTCTTGGTCAACCTATGTTGCTACTGCT 1500  
Db 1591 GGGGAACATGATCTCTGTTTCATCTCTGCTATTTTCGACAAACCGATGTTGCTGCTCT 1650  
Qy 1501 ATATACCATGATGATGATAGGAAGCAAACTTGCATGCTGAGCTACGG 1551  
Db 1651 TTATACACACCTGATGAACCGAAAGGATGATGCTGATGCTGAACCACTG 1701

## RESULT 7

AA51484

ID AA51484 standard; cDNA; 1985 BP.

XX

XX AA51484;

AC

DT 09-OCT-2000 (first entry)

DE

A. thaliana AS11 diacylglycerol acyltransferase cDNA.

XX

DGAT; diacylglycerol acyltransferase; seed oil; fatty acid synthesis;

XX

size; weight; carbon flux; TAG1; insertion mutant; ss.

XX

OS Arabidopsis thaliana.

XX

FH

FT

FT

FT

FT

FT

FT

FT

FT

XX

DN

XX

XX

PD

XX

XX

PF

XX

XX

PR

XX

XX

PA

XX

XX

PI

XX

XX

DR

XX

XX

PT

PT

PT

PT

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Location/Qualifiers

139..1782

/\*tag= a

/product= diacylglycerol\_acyltransferase

/note= "insertion mutant"

610..690

/\*tag= b

/note= "81 bp insertion, duplication of exon 2"

WO200036114-A1.

22-JUN-2000.

16-DEC-1999; 99WO-CA01202.

17-DEC-1998; 98US-0112812.

(CANA ) NAT RES COUNCIL CANADA.

Zou J, Taylor DC, Wei Y, Jako CC;

WPI; 2000-431592/37.

P-PSDB; AAY96854.

New DNA encoding diacylglycerol acyltransferase from Arabidopsis

thaliana for transforming plants and regulating seed oil content, fatty

acid synthesis and seed oil acyl composition in commercial and crop

plants

Claim 5; Page 79; 91pp; English.

This cDNA encodes Arabidopsis thaliana ecotype Columbia mutant AS11

diacylglycerol acyltransferase (DGAT). The AS11 mutant TAG1 allele has a

147 bp insertion located at the central region of intron 2. The insertion

is a duplication of a segment that is composed of 12 bp from the 3' end

of intron 1, the entire sequence of exon 2 (81 bp) and 54 bp from the

5' end of intron 2. The DGAT and the insertion mutant (AS11) are useful

for regulating seed oil content, the ratio of diacylglycerol to

triacylglycerol proportions in seed oil, fatty acid synthesis, seed oil

acyl composition, seed size/weight and carbon flux into other seed

components in commercial and crop plants. The natural formation of

triacylglycerols can be modified to increase the yield in commercial

plant oils or modify their composition to achieve specific commercial

improvements of plants and plant products.

Query Match 35.2%; Score 683; DB 21; Length 1985;

Best Local Similarity 74.4%; Pred. No. 1.7e-133;

Matches 898; Conservative 0; Mismatches 285; Indels 24; Gaps 2;

Qy 345 AGAGTCAGCGGGCTCTTCAACCTTGTATAGTAGTCTTGTGCTGTGAATAGCCGAC 404

Db 608 AGAGCCATGCGGATTTATCAACCTCTGTGTAGTAGTCTTATTGCTGTAACAGTAGAC 667

Qy 405 TCATCATTTGAGAATTTAATGAAGTATGTTGTTGATCAAAATCTGGCTTTGGTTAGTT 464

Db 668 TCATCATCGAAAATCTTATGAAGTATGTTGTTGATCAGAACGGATTTCTGGTTAGTT 727

Qy 465 CAAAGTCATGAGAGACTGGCCCTTTTCATGTTGTTGTTCTCTCTGTTGGTATTTCCTT 524

Db 728 CAAGATCGTCGAGATTTGGCGCTTTTCATGTTGTTGTTATATCCCTTTTCGATCTTTCCCTT 787

Qy 525 TCCTGCTTTTATAGTGGAGAGTTGGCACAACGGAAGTGTATACCCGAAACGATTGTTG 584

Db 788 TGGCTGCTTTTACGTTGAGAAATTTGGTACTTCAGAAATATACATATCAGAACCTGTTGCA 847

Qy 585 TTGTACTTCATATATATCACTTACCTCAACTTCTGCTTTTCTATCCAGTTTACTTATTTCA 644

Db 848 TCTTCTTCATATATATATCACCATGACAGAGTTTGTATCCAGTTTACCTACCCCTAA 907

Qy 645 GGTGTGATTCGCTTTTGTATATCAGGTGTCAGGTTAAATGCTGTTTCTTGTGTTGATGTT 704









Db	266	TTGTCCTTAGTCTGCCCTGCTTTCCCCCTGGGTGCATTTGCGATTTGAAAGTTGGCATTTAA	325
Qy	559	GAAGTGTATACCCGAACCAAGTGTGTGTGTACTATATATCATTAACCTCAACTTCGCT	618
Db	326	CAATGTTATTACTGATGCTGTTTGGCTACCTGCGCTCCATATCTTCTTCAACAACCGAAAT	385
Qy	619	TTTCTATCCAGTTTATGTTATCTTCAGGTGTGATTCCTGTTTTTGTATCAGGTGTCAGTT	678
Db	386	TGTATATCCAGCTGTTGTGATTTCTTAAGTGTGATTCGACGTTTGTCTGGCTTTTGT	445
Qy	679	AATGCTGTTTCTTGTGTTGTATGCTTAAATTTGGTGCTTTATGCACATACAAACTATGA	738
Db	446	GATATTATGCTGTATTTGTTGGCTGAGCGTTGTATCTTTTGCACATACAAACCATGA	505
Qy	739	TATGAGAGCACTTACCAAAATTAGTTGAAAGGGAGAAGCACTGCTCGATACTCTGAACAT	798
Db	506	TATAAGCAACTGACCATGGCGGCAAGAAGGTTGATTAATGAACATAAGCACAGTTGACAT	565
Qy	799	GGACTATCCTTACAACGTAAAGCTTCNAAGAGCTTGGCATATTTCTGTTGCCCTACATT	858
Db	566	GGATAATTACAAACCTTCAACACTTTAGGGAATCTAATACTTCAATGATGGCTCCTACACT	625
Qy	859	ATGTTACCAAGCAAGTATCCTCGCACACCTTATATTCGAAAGGGTGTGTTTTCGCCA	918
Db	626	CTGTTATCAGCAAGCATATCCCGRACITCATGTGTTAGAAAAGTTGGCTGATTCGTCA	685
Qy	919	ACTTGTCAAGCTGATTAATATTTTACAGGAGTTATGCGGATTTATATAGACCAATATATA	978
Db	686	AATTTATCTGTACTTGTATCTTACTGGCTTCAAGGCCTCATATTATGAGCAATACATAAA	745
Qy	979	TCCATAGTACAAAATTCACAGCATCCTCTCAAGGAAACCTTCTTAGCCACCCGAGAG	1038
Db	746	TCCANTGTTGCAATTCACGATCCCATGGAAGGAGCACTCCTAAATGCTGTAGAGAC	805
Qy	1039	AGTCTGAAGCTTCTGTTCCTTCCAAATTTATGTGTGGCTTCGATGTTCTATTGCTTTTT	1098
Db	806	TGTTTTGAAACTCTCATACCAATGTTTACCTGGCTTTTGCATGTCTATGCTTTTTT	865
Qy	1099	CCACCTTGGTTAAATATCCTGGCAGAGCTTCTCGATTTGGTGATCGTGAATCTCAAA	1158
Db	866	CCATCTCTGGTTAAGTATACTTGCCTGAGATCTTCGATTTGGTGACCGTGAATTTCTACA	925
Qy	1159	GGATTGTGGAATGCCAAACTGTCCGAGATATTGGAGGATGGAATATGCTGTGTCA	1218
Db	926	AGATTGGTGGATGCAAAAACAATTTGATGAGTATTGGAGAAAATGGAATATGCTGTACA	985
Qy	1219	CAATGGATGATCCGCCCACTATATTTTCCATGTTTTAAGGCACCGCTACCAAGGCTGC	1278
Db	986	TAAATGGGTGTTGCGCATATTACTTTCTTGCATGCGAAATGGTATATCAAGGAAGT	1045
Qy	1279	TGCTCTTTTAATTCCTCCTGGTTCTGCTTTATTCCTCAGCTGCCATTCGTGTTC	1338
Db	1046	TGCTGTCTGATATCATCTTGTGTTCTGCGCGTACTCCTCAGATATGTGTGCGTGTTC	1105
Qy	1339	TTGCCACATATTCAGTTTGGGCTTTCGGTGGAAATATGTTTTCAGGTTCCTTTGGTCTT	1398
Db	1106	CTGCCGATTCACAAGTTCTGGGCATCTCTTAGGAATATGCTACAGATCCCCCTTATCGT	1165
Qy	1399	GATCACTAATTATCTGCAAAATTAATATTCAGAAACTCAATGTTTGGAAATATGATTTTTG	1458
Db	1166	ATTGACAGCATACCTCAAAAGTAAATTCAGAGATACAAATGGTTGGCAACATGATATTTG	1225
Qy	1459	GTTCTATATCAGTATCTTGGTCAACCTATGTGTCTACTGCTACTACTACCATGACTTGAT	1518
Db	1226	GTTCTTTTCTGCATCTATGGCGCAATGTGCCCTCTCCTGTACTATCATGATGTGAT	1285
Qy	1519	GAATAGGAAGGCCAA	1533
Db	1286	GAACGAGATTGAGAA	1300

RESULT 10  
AAA48933

AA48933 standard; cDNA; 1281 BP.	
AAA48933;	
06-DEC-2000 (first entry)	
Corn diacylglycerol acyltransferase cDNA #1.	
Diacylglycerol acyltransferase; corn; rice; soybean; wheat;	
triacylglycerol; herbicide; EC2.3.1.20; ss.	
Zea mays.	
Key Location/Qualifiers	
CDS 1..1090	
FT /*tag= a	
FT /partial	
FT /transl_except= (pos:283..286,aa:Ile)	
FT /note= "This codon has an apparent 1 nucleotide	
FT insertion which alters the reading frame"	
FT /transl_except= (pos:356..358,aa:Xaa)	
FT /note= "Xaa= unknown"	
FT /transl_except= (pos:869..889,aa:YVLLFL)	
FT /product= Partial_diacylglycerol_acyltransferase	
PN W0200032756-A2.	
XX 08-JUN-2000.	
XX 01-DEC-1999; 99WO-US28354.	
XX 02-DEC-1998; 98US-0110602.	
XX 31-MAR-1999; 99US-0127111.	
XX (DUPO ) DU PONT DE NEMOURS & CO E I.	
XX Cahoon EB, Kinney AJ, Cahoon RE;	
XX WPI; 2000-412308/35.	
XX P-PSDB; AAY94513.	
XX Polynucleotides encoding diacylglycerol acetyltransferase, useful for	
XX synthesis of triacylglycerols and increasing the level of oils in plant	
XX seeds .	
XX Claim 15; Page 40-42; 62pp; English.	
XX In the present invention, cDNA libraries from Arabidopsis, corn, rice,	
XX soybean and wheat were screened for sequences with homology to a	
XX putative acyl CoA cholesterol acyltransferase related gene from	
XX Arabidopsis thaliana and diacylglycerol acyltransferases from Homo	
XX sapiens and Mus musculus. The cDNA clones identified from this process	
XX were used to form complete diacylglycerol acyltransferase cDNA	
XX sequences. The present sequence is corn diacylglycerol acyltransferase	
XX cDNA from a contig of clones cpj1c.pk005.h23, cen3n.pk0010.c10,	
XX cc01.pk0029.b6. Diacylglycerol acyltransferases are involved in the	
XX synthesis of triacylglycerols. Alteration of the expression of the	
XX diacylglycerol acyltransferase DNA can be useful for increasing the	
XX level of oils in plant seeds. Inhibitors of diacylglycerol	
XX acyltransferase may be useful as herbicides.	
XX Sequence 1281 BP; 351 A; 256 C; 245 G; 417 T; 12 other;	

	Query Match	24.7%;	Score 478.8;	DB 21;	Length 1281;
	Best Local Similarity	66.4%;	pred. No. 8.7e-91;		
	Matches 717;	Conservative 0;	Mismatches 359;	Indels 4;	Gaps 2;
-Qy.	458	TTTAGTTCAAAGTCATTGAGACACTGGCCCCCTTTTCATGTGTTGCTCTTCCTTGTGGTA	517		
Db	1	TTTAATGTCATCATCTCGAGACTGGCCACTGCTTAATGTGTGGCTTAGTCTACCCATA	60		
-Qy	518	TTTCCTTTTCGGTGGCTTTTATAGTGAGAGAAGTTGGCACACACGGAAGTGATACCCGAACCA	577		









QY 954 GATTTATTAATAGACCAATATATTAATCCCATAGTACAAAAATTCACAGCATCCTCTCAAGG 1013  
Db 70 GATTCATATTGAGCAATACATAATCCTATTGTGTGAAGCTCTCAACATCCATTGATGG 129  
QY 1014 GAAACCTTCTTTAGCCACCGAGAGAGTCTGAAAGCTTTCTGTCCAAATTTATATGTGT 1073  
Db 130 GAGGATTACTGAATGCTGTAGAGACTGTTTTGAAGCTCTCATTAACCAATGCTACCTGT 189  
QY 1074 GGCCTGCGATGTTCTATTGCTTTTTCCACCTTTTGGTTAAATATCCTGGCAGAGCTTCTTC 1133  
Db 190 GGCCTTGCATGTTTATTGCTTTTTCCATCTGTGTTAAACATACTTGCAGATTCTTC 249  
QY 1134 GATTTGGTGATCGTGAATCTACAAAGGATTTGGTGGAAATGCCAAACTGTCGAAGATTATT 1193  
Db 250 GATTTGGTGACCGAGAATCTTACAAAGACTGGTGGAAATGCAAGACAATTGATGAGTACT 309  
QY 1194 GGAGGATGTGGAATATGCTGTTTACAAATGGATGATCCGCCACCTATATTTCCATGTT 1253  
Db 310 GGAGAAATGGAACATGCTGTGCATAAATGGATTGTTGTCATATATATTTCCCTTGCA 369  
QY 1254 TAAGGCACGGTCTACCAAGGCTGCTGCTCTTTTAATTGCTTCTGCTGCTTTAT 1313  
Db 370 TCGCAATGGTATATCAAGGAAGTGTCTGTTTTTATATCGTTCTTCTGTTCTGCTGTAC 429  
QY 1314 TCCATGAGCTGTGCATTGCT 1333  
Db 430 TTCATGAGCTGCAGATTACT 449

Search completed: February 21, 2003, 16:23:06  
Job time : 448 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 16:13:34 ; Search time 2523 Seconds  
(without alignments)  
12465.967 Million cell updates/sec

Title: US-09-856-018b-15  
Perfect score: 1942  
Sequence: 1 tagaacaacagctcgtctt.....aaaaaaaaaaaaaaaaaaaaa 1942

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_mam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	816.6	42.0	855	10 AW349274	AW349274 GM210004B
c 2	488	25.1	961	14 BQ510367	BQ510367 EST617782
c 3	470.2	24.2	572	13 BM309699	BM309699 sak65f03
4	461.4	23.8	586	14 BQ148998	BQ148998 NF086D09F
5	433	22.3	433	9 A1441040	A1441040 sa58f02.y
6	432.2	22.3	729	13 B1422326	B1422326 EST532992

7	427.6	22.0	1181	11 AY110660	AY110660 zea mays
8	424.6	21.9	572	14 BQ124305	BQ124305 EST609881
9	404.8	20.8	1572	11 AY105372	AY105372 zea mays
10	368.2	19.0	741	14 BQ998922	BQ998922 QG20J23
11	334.6	17.2	459	14 BQ123670	BQ123670 EST609246
12	321.8	16.6	539	13 B1422212	B1422212 EST532878
13	307.4	15.8	682	10 BE247899	BE247899 NF038D11D
14	305	15.7	763	13 B1308446	B1308446 EST529856
15	295.8	15.2	729	13 B1322102	B1322102 BJ322102
16	295	15.2	577	10 AW586836	AW586836 EST318459
17	290.6	15.0	634	12 BF634363	BF634363 NF059D06D
18	288	14.8	540	9 AJ470192	AJ470192 AJ470192
19	282.8	14.6	606	13 B1316561	B1316561
20	278.8	14.3	626	10 AV926912	AV926912 AV926912
21	276.8	14.3	561	10 AW035727	AW035727 EST281881
22	255.6	13.2	655	14 BQ998311	BQ998311 QG19C15
23	254.4	13.1	862	12 BG321213	BG321213 Zm04_05g0
24	245.4	12.6	480	9 AJ470191	AJ470191 AJ470191
25	239.6	12.3	480	9 AJ470190	AJ470190 AJ470190
26	238	12.3	528	9 AL381190	AL381190 MCB57D12
27	237	12.0	508	14 BU009402	BU009402 QGJ10F20
28	233.8	12.0	673	14 BQ998554	BQ998554 QG19N20
29	229.8	11.8	629	9 AA042298	AA042298 24635 CD4
30	220.6	11.4	515	10 AV925760	AV925760 AV925760
31	217.4	11.2	555	13 B1268713	B1268713 B1268713
32	216.8	11.2	685	14 BQ861203	BQ861203 QG17N02
33	212.4	10.9	656	14 BQ862755	BQ862755 QG21P16
34	200.4	10.3	422	12 BF199515	BF199515 WHE0591-0
35	198.6	10.2	617	10 AV930897	AV930897 AV930897
36	180	9.3	531	10 BE356202	BE356202 DG1_123_B
37	178.8	9.2	238	12 BF066952	BF066952 st35e10.y
38	178.6	9.2	706	13 BM348355	BM348355 MEST289-G
39	176	9.1	655	10 AW775077	AW775077 EST334228
40	174.4	9.0	574	12 BF003479	BF003479 EST431977
41	170.8	8.8	701	13 BM334936	BM334936 MEST130-B
42	168.2	8.7	700	13 BM341343	BM341343 MEST133-F
43	164.2	8.5	682	13 BM075161	BM075161 MEST351-B
44	162.6	8.4	681	13 BM267747	BM267747 MEST371-E
45	161.6	8.3	627	10 AV956927	AV956927 AV956927

ALIGNMENTS

RESULT 1  
AW349274/c

LOCUS	AW349274	855 bp	mRNA	linear	EST 04-OCT-2000
DEFINITION	GM210004B21H12 Gm-r1021 Glycine max CDNA clone Gm-r1021-1536 3',				
ACCESSION	AW349274				
VERSION	AW349274.1				
KEYWORDS	EST.				
SOURCE	soybean.				
ORGANISM	Glycine max				
REFERENCE	1 (bases 1 to 855)				
AUTHORS	Vodkin,L., Kelm,P., Shoemaker,R., Retzel,E., Khanna,A., Corvelli,V., Erpelting,J., Raph,C., Shoop,E., Pardinas,J., Liu,L. and Lewin,H.				
TITLE	A Functional Genomics Program for Soybean (NSF 9872565)				
JOURNAL	Unpublished (1999)				
COMMENT	Other ESTs: A1441040 Contact: Vodkin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565) Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics University of Illinois Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA Tel: (217) 244-6147 Fax: (217) 333-4582 Email: l-vodkin@uiuc.edu				

	FEATURES	SOURCE
1.	100% Pure Cotton	USA
2.	Machine Washable	India
3.	Soft & Comfortable	Bangladesh
4.	Durable & Long Lasting	Pakistan
5.	Eco-Friendly	Vietnam
6.	Available in Multiple Colors	Thailand
7.	Lightweight & Breathable	Philippines
8.	Easy to Fold & Store	Malaysia
9.	Hypoallergenic	Singapore
10.	Perfect for Travel	Japan

Query Match 42.0%; Score 816.6; DB 10; Length 855;

Matches	819	Conservative	0	Mismatches	35	Indels	0	Gaps	0
QY	1010	AAGGGAAACCTCTTTACGCCACCGAGAGAGTCTTGGAAGCTTCTGTGTCCAAATTTATAT	1069						
Db	854	AAGGAAACCTTNNNTTACGCCACNNNNAGTNNNGAANCNNNNGTTCCAAATTTATAT	795						
QY	1070	GTGTGGCTGTGCATGTCTATTCGCTTTTCCACCTTTTGGTTAAATATCCTGGGCAGACGTT	1129						
Db	794	GTGTGNCCTGTGCATGTCTATTCGCTTTTCCNNNNNNNGNTAAATATCCTGGGCAGACGTT	735						
QY	1130	CTTTCGATTTGGTGATCGTGAATTTCTACAAGGATTTGGTGAATGCCAAACTGTGCGAAGAT	1189						
Db	734	CTTTCGATTTGGTGATCGTGAATTTCTACAAGGATTTGGTGAATGCCAAACTGTGCGAAGAT	675						
QY	1190	TATTTGGAGGATGTGGAATATGCGCTGTTCACAAATGGATGATCCGCCACCTATATTTTCCA	1249						
Db	674	TATTTGGAGGATGTGGAATATGCGCTGTTCACAAATGGATGATCCGCCACCTATATTTTCCA	615						
QY	1250	TGTTTAAAGGCACGGCTCTACCAAAGGCTGTGCTCTTTTAAATTTGGCTTCCTGGTTCTCGCT	1309						
Db	614	TGTTTAAAGGCACGGCTCTACCAAAGGCTGTGCTCTTTTAAATTTGGNNNNCTGGTTCTCGCT	555						
QY	1310	TTATTTCCATGAGCTGTGCATTTGCTGTTCCTTGGCCACATATTCGAATGTGTGGGCTTCGGT	1369						
Db	554	TTATTTCCATGAGCTGTGCATTTGCTGTTCCTTGGCCACATATTCGAATGTGTGGGCTTCGGT	495						
QY	1370	GGAAATATGTTTTCAGGTTCCTTTTGGTCTTGATCACTAAATTTATCTGCAGAAATAAATTCAGA	1429						
Db	494	GGAAATATGTTTTCAGGTTCCTTTTGGTCTTGATCACTAAATTTATCTGCAGAAATAAATTCAGA	435						
QY	1430	AACTCAATGGTTGGAATAATGATTTTTTGGTTCATATTCAGTATCCTTTGGTTCACACCTATG	1489						
Db	434	AACTCAATGGTTGGAATAATGATTTTTTGGTTCATATTCAGTATCCTTTGGTTCACACCTATG	375						

QY	1490	TGTGTA	CTGCTAT	CTACTAC	CAATG	ACTT	GATGA	TAGGA	AGGCA	AAC	TTG	ACT	GAAG	CTAC	1549
Db	374	TGTGTA <th>CTGCTAT</th> <th>CTACTAC</th> <th>CAATG</th> <th>ACTT</th> <th>GATGA</th> <th>TAGGA</th> <th>AGGCA</th> <th>AAC</th> <th>TTG</th> <th>ACT</th> <th>GAAG</th> <th>CTAC</th> <td>315</td>	CTGCTAT	CTACTAC	CAATG	ACTT	GATGA	TAGGA	AGGCA	AAC	TTG	ACT	GAAG	CTAC	315
QY	1550	GGCATT <th>TACAT</th> <th>TTTAA</th> <th>AGGTG</th> <th>GCACAT</th> <th>GGATG</th> <th>AGCTTT</th> <th>TCAGTT</th> <th>TTTCAG</th> <th>TTTTCAG</th> <th>TTTTCAG</th> <th>TTTTCAG</th> <th>TTTTCAG</th> <td>1609</td>	TACAT	TTTAA	AGGTG	GCACAT	GGATG	AGCTTT	TCAGTT	TTTCAG	TTTTCAG	TTTTCAG	TTTTCAG	TTTTCAG	1609
Db	314	GGCATT <th>TACAT</th> <th>TTTAA</th> <th>AGGTG</th> <th>GCACAT</th> <th>GGATG</th> <th>AGCTTT</th> <th>TCAGTT</th> <th>TTTTCAG</th> <th>TTTTCAG</th> <th>TTTTCAG</th> <th>TTTTCAG</th> <th>TTTTCAG</th> <td>255</td>	TACAT	TTTAA	AGGTG	GCACAT	GGATG	AGCTTT	TCAGTT	TTTTCAG	TTTTCAG	TTTTCAG	TTTTCAG	TTTTCAG	255
QY	1610	ATGTGA <th>TATG</th> <th>TTGGT</th> <th>CAATAT</th> <th>TTGTTT</th> <th>CTAG</th> <th>GAATG</th> <th>CTTTT</th> <th>TCATCT</th> <th>PACCAT</th> <th>GCGCAT</th> <th>TGG</th> <td>1669</td>	TATG	TTGGT	CAATAT	TTGTTT	CTAG	GAATG	CTTTT	TCATCT	PACCAT	GCGCAT	TGG	1669	
Db	254	ATGTGA <th>TATG</th> <th>TTGGT</th> <th>CAATAT</th> <th>TTGTTT</th> <th>CTAG</th> <th>GAATG</th> <th>CTTTT</th> <th>TCATCT</th> <th>PACCAT</th> <th>GCGCAT</th> <th>TGG</th> <td>195</td>	TATG	TTGGT	CAATAT	TTGTTT	CTAG	GAATG	CTTTT	TCATCT	PACCAT	GCGCAT	TGG	195	
QY	1670	CTGCTC <th>TGGA</th> <th>AGGA</th> <th>ATCC</th> <th>ACGG</th> <th>ATATG</th> <th>CCAG</th> <th>ATATG</th> <th>CCAG</th> <th>ATATG</th> <th>CCAG</th> <th>ATATG</th> <td>1729</td>	TGGA	AGGA	ATCC	ACGG	ATATG	CCAG	ATATG	CCAG	ATATG	CCAG	ATATG	1729	
Db	194	CTGCTC <th>TGGA</th> <th>AGGA</th> <th>ATCC</th> <th>ACGG</th> <th>ATATG</th> <th>CCAG</th> <th>ATATG</th> <th>CCAG</th> <th>ATATG</th> <th>CCAG</th> <th>ATATG</th> <td>135</td>	TGGA	AGGA	ATCC	ACGG	ATATG	CCAG	ATATG	CCAG	ATATG	CCAG	ATATG	135	
QY	1730	TATGTACT <th>TACCA</th> <th>CTCT</th> <th>CTCT</th> <th>CGCA</th> <th>ATG</th> <th>TATCA</th> <th>AAATAT</th> <th>GCAAT</th> <th>TTTTCAG</th> <th>AGG</th> <th>CCAT</th> <td>1789</td>	TACCA	CTCT	CTCT	CGCA	ATG	TATCA	AAATAT	GCAAT	TTTTCAG	AGG	CCAT	1789	
Db	134	TATGTACT <th>TACCA</th> <th>CTCT</th> <th>CTCT</th> <th>CGCA</th> <th>ATG</th> <th>TATCA</th> <th>AAATAT</th> <th>GCAAT</th> <th>TTTTCAG</th> <th>AGG</th> <th>CCAT</th> <td>75</td>	TACCA	CTCT	CTCT	CGCA	ATG	TATCA	AAATAT	GCAAT	TTTTCAG	AGG	CCAT	75	
QY	1790	CAC <th>TGGC</th> <th>ATTG</th> <th>ATAA</th> <th>CTG</th> <th>CCAG</th> <th>GAAC</th> <th>CACT</th> <th>CTAA</th> <th>CTCT</th> <th>TTTCT</th> <th>CTG</th> <th>TTTAA</th> <td>1849</td>	TGGC	ATTG	ATAA	CTG	CCAG	GAAC	CACT	CTAA	CTCT	TTTCT	CTG	TTTAA	1849
Db	74	CAC <th>TGGC</th> <th>ATTG</th> <th>ATAA</th> <th>CTG</th> <th>CCAG</th> <th>GAAC</th> <th>CACT</th> <th>CTAA</th> <th>CTCT</th> <th>TTTCT</th> <th>CTG</th> <th>TTTAA</th> <td>15</td>	TGGC	ATTG	ATAA	CTG	CCAG	GAAC	CACT	CTAA	CTCT	TTTCT	CTG	TTTAA	15
QY	1850	AGAGG <th>CGCTAG</th> <th>ATGT</th> <td>1863</td> <td data-cs="10" data-kind="parent"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td>	CGCTAG	ATGT	1863										
Db	14	AGAGG <th>CGCTAG</th> <th>ATGT</th> <td>1</td> <td data-cs="10" data-kind="parent"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td>	CGCTAG	ATGT	1										
RESULT 2															
BQ510367/c															
LOCUS															
DEFINITION															
BQ510367 961 bp mRNA linear EST 22-JUL-2002															
EST617782 Generation of a set of potato cDNA clones for microarray															
analyses mixed potato tissues Solanum tuberosum cDNA clone STMHJ76															
3' end, mRNA sequence.															
ACCESSION															
VERSION															
KEYWORDS															
SOURCE															
ORGANISM															
Solanum tuberosum															
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;															
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;															



axillary buds of stem explants, petioles, germinating eyes	
BASE COUNT	324 a 205 c 181 g 251 t
ORIGIN	
Query Match	25.1%; Score 488; DB 14; Length 961;
Best Local Similarity	75.2%; Pred. No. 5.3e-68;
Matches 627; Conservative 0; Mismatches 195; Indels 12; Gaps 1;	
QY 716	TCTATGACATACAACTATGATGAGAGCACTTACCATAATAGTTGAAAAGGAGAA 775
DB	
QY 961	TCTATGACATACAAATATGATGAGAGCACTTGCAGAGTCTGGATGAGGTGAA 902
DB	
QY 776	GCAGTCTCGATCTGACATGAGTATCTTACAGCTTAAGCTTCAAGAGCTTGGCA 835
DB	
QY 901	-----AATTCGGAATCACTACTCTTACAATGTTAGTTTCAAGAGTTTGGCT 854
DB	
QY 836	TATTTCTGTTGCTCCCTACATTTATGTTACAGCAAGTATCTCTCGCACCTTATAT 895
DB	
QY 853	TACTTCTATGTTGCTCCAACTTTATGCTATGCTATGCTATGCTCTGCTGCTCAT 794
DB	
QY 896	CGAAGGGTGTGTTTTCGCAACTTGTCAAGCTGATATATTTACAGAGTTATGGGA 955
DB	
QY 793	CGAAGGGTGTGTTTTCGCAACTTGTCAAGCTGATATATTTACAGAGTTATGGGA 734
DB	
QY 956	TTTATATACCAATATATTAATCCATAGTACAAAATTCACAGCATCTCTCAAGGGA 1015
DB	
QY 733	TTTATCATTCAGCAGTATATTAACCGGATTTGCGAAGCTCACAACATCCATTTGA 674
DB	
QY 1016	AACCTTCTTTACGCCACCGAGAGTTCTGAAGCTTCTGTTCCAAATTTATATGTTGG 1075
DB	
QY 673	AACCTTCTTTACGCCACCGAGAGTTCTGAAGCTTCTGTTCCAAATTTATATGTTGG 614
DB	
QY 1076	CTCTGATGTTCTATGCTTTTTCACACCTTGGTTAAATATCTCTGCGAGAGTCTTTCGA 1135
DB	
QY 613	CTCTGATGTTCTATGCTTTTTCACACCTTGGTTAAATATCTCTGCGAGAGTCTTTCGA 554
DB	
QY 1136	TTTGGTATCGTGAATTCACAGGATTTGGTGAATGCGCAAACTCTCAAGATTTATGG 1195
DB	
QY 553	TTTGGGATCGTGAATTCACAGGATTTGGTGAATGCGCAAACTCTCAAGATTTATGG 494
DB	
QY 1196	AGGATGGAATATGCTTGTTCACAAATGGATGATCGCGCACTATATTTTCCATGTTA 1255
DB	
QY 493	AGATTTGGAATATGCTTGTTCACAAATGGATGATCGCGCACTATATTTTCCATGTTA 434
DB	
QY 1256	AGGACGGTCTACCAAGGCTGCTCTTTAAATGCTTCTCTGTTTCTGCTTTTATTC 1315
DB	
QY 433	AGGATGGAATATGCTTGTTCACAAATGGATGATCGCGCACTATATTTTCCATGTTA 374
DB	
QY 1316	CATGAGCTGATGCTGCTTCTTCCACATATTCAGTTTGGGCTTTCGGTGGAAAT 1375
DB	
QY 373	CATGAGCTGATGCTGCTTCTTCCACATATTCAGTTTGGGCTTTCGGTGGAAAT 314
DB	
QY 1376	ATGTTTTCAGGTTCTTGTGTTGATCACTAATATTCGCAAAATAAATTCAGAACTCA 1435
DB	
QY 313	ATGTTTTCAGGTTCTTGTGTTGATCACTAATATTCGCAAAATAAATTCAGAACTCG 254
DB	
QY 1436	ATGTTTTCAGGTTCTTGTGTTGATCACTAATATTCGCAAAATAAATTCAGAACTCG 1495
DB	
QY 253	AATGTGGGCAACATGACATTCCTGTTGTTTCTGCTGTTTGGTCAACCAATGTTGTG 194
DB	
QY 1496	CTGCTATATCATGACATTTGATGAATAGGAAGGCAAACTTCAGTGAAGCTTAC 1549
DB	
QY 193	CTTCTGTTATACCATGATGATGAATAGGAATGTTAGTTTAAAGCTTTC 140
DB	
RESULT 3	
BM309699	
LOCUS	572 bp mRNA linear EST 02-JAN-2002
DEFINITION	sak65f03.y1 Gm-c1036 Glycine max cDNA clone SOYBEAN CLONE ID:
DIACYLGLYCEROL ACYLTRANSFERASE ;, mRNA sequence.	
ACCESSION	BM309699

BM309699.1	GI:18041405
EST.	
soybean.	
Glycine max	
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;	
Glycine.	
1 (bases 1 to 572)	
Shoemaker,R., Keim,P., Vodkin,L., Erpellding,J., Coryell,V., Khanna	
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,	
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,	
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk	
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann	
,R., Waterston,R. and Wilson,R.	
Public Soybean EST Project	
Unpublished (1999)	
Contact: Shoemaker R/Public Soybean EST Project	
Public Soybean EST Project	
Washington University School of Medicine	
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA	
Tel: 314 286 1800	
Fax: 314 286 1810	
Email: est@watson.wustl.edu	
This clone is available through: Resgen, Invitrogen Corp. 2130	
South Memorial Parkway Huntsville, AL 35801 For further information	
call: (800)-533-4363 or contact: ccu@resgen.com web site:	
www.resgen.com	
Seq primer: -40RP from Gibco	
High quality sequence stop: 412.	
Location/Qualifiers	
1..572	
/organism="Glycine max"	
/db_xref="taxon:3847"	
/clone="SOYBEAN CLONE ID: Gm-c1036-7949"	
/clone_lib="Gm-cl036"	
/tissue_type="Somatic embryos cultured on MSD 20"	
/lab_host="DH10B"	
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; This	
cDNA library was constructed from mRNA isolated from	
somatic embryos (age ranging from 2 months to 9 months)	
cultured on MSD 20. The library was prepared using the	
Life Technologies pSuperScript cDNA library construction	
kit. Complementary DNA was synthesized from mRNA using a	
poly (dT) sequence with a NotI restriction site. SalI	
linkers adapters were ligated to the blunt-ended cDNA	
fragments followed by NotI digestion. The cDNA fragments	
were directionally cloned into the NotI-SalI restriction	
site of the pSPORT1 vector. The ligated cDNA fragments	
were transformed into E.coli Electromax DH10B host cells.	
This library was constructed in the laboratory of Dr. Lila	
Vodkin by Anu Khanna at the University of Illinois at	
Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"	
BASE COUNT	124 a 169 c 117 g 161 t
ORIGIN	
Query Match	24.2%; Score 470.2; DB 13; Length 572;
Best Local Similarity	92.7%; Pred. No. 4.4e-65;
Matches 532; Conservative 0; Mismatches 24; Indels 18; Gaps 3;	
QY 55	TGTAGGCACTGCTCTCAACCACTCTTCCCTGCGCGCGCTCCCTCCGCACTCCACCGC 114
DB	
QY 17	TGTACCACTGCTCTCAACCACTCTTCCCTGCGCGCGCTCCACCGC-----CCGC 67
DB	
QY 115	CGGCTCTTCAATTCGCTGAGACAACCGACAGTTCGGTGATGACTTGCCCAAGGA 174
DB	
QY 68	TGGCTCTTCAATTCGCGAGACGACCGACAGTTCGGTGATGACTTGCCCAAGGA 127
DB	
QY 175	TTCTGTTCCGAGCACTTCCATCAACGACGACGCGCGCTCAATTCGCAACGACAAA 234
DB	
QY 128	TTCCGTTCCGAGCACTTCCATCAGACGCA-----CGCGCAATTCGCAACGCA-- 179
DB	
QY 235	CGAAAAAAGACACATGATTTCTCCGCTCTCTCAAAATTCGGCTACCGCTTCCGCTCCCGC 294

Db 180 -CAAAACAGACACTGATTTTCGGTCTCAAAATTCGCTACCGTCTCCGTCGCCGC 238  
 Qy 295 TCACCGCAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTTCCGTGACAGATCAGCG 354  
 Db 239 TCATCGCAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTTCCGTGACAGATCAGCG 298  
 Qy 355 GGGCTCTTCAACCTTTGTATAGTAGTCCCTTGTGCTGTGAATAGCCGACTCATCTGA 414  
 Db 299 GGGCTCTTCAACCTCTCTATAGTAGTCCCTTGTGCTGTGAATAGCCGACTCATCTGA 358  
 Qy 415 GAAATTAATGAATAGTGTGGTGAATCAAAATCAGCTTTGGTTAGTTCAAGTCAATT 474  
 Db 359 GAAATTAATGAATAGTGTGGTGAATCAAAATCAGCTTTGGTTAGTTCAAGTCAATT 418  
 Qy 475 GAGACATGGCCCTTTTCATGTGTGCTTCTCTGCTGTATTTCCCTTTCGCTGCCCTT 534  
 Db 419 GAGACATGGCCCTCTTCATGTGTGCTTCTCTGCTGTATTTCCCTTTCGCTGCCCTT 478  
 Qy 535 TATAGTGAGAAAGTGGCACACGGAAGTGTATACCCGAAACAGTGTGTGTGACTTCA 594  
 Db 479 TATAGTGAGAAAGTGGCACACGGAAGTGTATACCCGAAACAGTGTGTGTGACTTCA 538  
 Qy 595 TATATCATTTACCTCAACTTCGCTTTTCTATCCA 628  
 Db 539 TATATCATTTACCTCAGCTTCACTTTTCTATCCA 572

## RESULT 4

B0148998 586 bp mRNA linear EST 24-APR-2002  
 LOCUS NF086D09FL1F1077 Developing flower Medicago truncatula cDNA clone  
 DEFINITION NF086D09FL 5', mRNA sequence.

ACCESSION B0148998  
 VERSION B0148998.1 GI:20286057

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

REFERENCE 1 (bases 1 to 586)

Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,  
 Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula flower library

Unpublished (2001)

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7391

Fax: 580 221 7380

Email: gdmay@noble.org

Insert Length: 586 Std Error: 0.00

Plate: 086 row: D column: 09

Seq primer: TCACACAGAAACAGCTATGAC.

Location/Qualifiers

FEATURES

source

1..586

/organism="Medicago truncatula"

/db\_xref="taxon:3880"

/clone="NF086D09FL"

/tissue\_type="Developing flower"

/dev\_stage="Developmentally pooled"

very young, developing, fully-opened flowers and flowers

in early transition into pods."

/note="vector: Lambda Zap; cDNA was prepared from polyA+

enriched, pooled samples of equivalent amounts of total

RNA from very young, developing, fully-opened flowers and

flowers transitioning into pods. The cDNA was

directionally ligated into the Uni-Zap XR vector

(Stratagene) and packaged using the Gigapack III Gold  
 packaging extracts. Phagemids containing cDNA inserts were  
 in vivo excised from the recombinant Uni-Zap XR vector  
 using EXAssist helper phage and the E. coli strain  
 XLI-Blue MRF (Stratagene). Excised plasmids were plated  
 using SOLR cells."

BASE COUNT 134 a 115 c 132 g 202 t 3 others  
 ORIGIN

Query Match 23.8%; Score 461.4; DB 14; Length 586;

Best Local Similarity 86.5%; Pred. No. 1.1e-63;

Matches 507; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 837 ATTTCCTGGTGGCCCTACATTATGTTACAGCCAAAGCTATCTCGCACACCTTATATTC 896

Db 1 ACTTCATGGTGGCTTACATTATGCTACAGCCAAAGCTATCTCGCACACCTTCGGTTC 60

Qy 897 GAAAGGTTGGTGTGGCCCAACTTGTCAAGCTGATATATTTACAGAGAGTTATGGAT 956

Db 61 GAAAGGTTGGTGTGGCCCAACTTGTCAAGCTGATATATTTACAGAGAGTTATGGAT 120

Qy 957 TTATAATAGACCAATATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGAA 1016

Db 121 TTATAATAGACCAATATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGAA 180

Qy 1017 ACCTTCTTTACGCCACCGAGAGAGTTCTGAAGCTTTCTGTTCCAAATTTATATGTGTGC 1076

Db 181 ACCTTCTATATGCCATTGAGAGAGTTCTGAAGCTTTCTGTTCCAAATTTATGTGTGC 240

Qy 1077 TCTGCATGTTCTATGCTTTTCCACCTTTGGTTAAATATATCTCGCAGAGTTCTTCGAT 1136

Db 241 TGTGCATGTTCTATGCTTTTCCATCTTTGGTTAAATATATCTCGCAGAGTTCTTCGAT 300

Qy 1137 TTGTGTGATCGTGAATTTCTACAGGATTTGGTGAATGTCGCAAACTGTCGCAAGATTTCGA 1196

Db 301 TTGTGTGATCGTGAATTTCTACAGGATTTGGTGAATGTCGCAAACTGTCGCAAGATTTCGA 360

Qy 1197 GGATGTGGAATATGCCTGTTTCCAAATGGATGATCCGCCACCTATATTTCCATGTTTAA 1256

Db 361 GGATGTGGAATATGCCTGTTTCCAAATGGATGATGTCGACGCTATTTTCCCTGTCATAA 420

Qy 1257 GGCAGGCTACCAAAAGCTGCTGCTTTTAAATGGCTTCTCGTGTCTCTGCTTTATTC 1316

Db 421 GGTGTGATATACCAAGGGTCTGCTGTTTAAATGGCTTCTCGTGTCTCTGCTTTATTC 480

Qy 1317 ATGAGCTGTGCATGCTGTTCTTCCACATATTCAGTGTGGCTTTCGTCGTAATTA 1376

Db 481 ATGAGTATGCAATGCTGTTCTTCCCAATGTCGCTTTCGCTTTTATTTGGAATTA 540

Qy 1377 TGTTTCAGGTTCTTTGGTCTTGTATCATTAATATCTGCAAAATAA 1422

Db 541 TGTTCCANGNTCTTTGGTCTTGTATCACCATTACCTGAAAAATAA 586

## RESULT 5

AI441040

LOCUS

DEFINITION

AI441040

VERSION

AI441040.1

KEYWORDS

EST.

SOURCE

soybean.

Glycine max

ORGANISM

Eukaryota;

Viridiplantae;

Streptophyta;

Embryophyta;

Tracheophyta;

Spermatophyta;

Magnoliophyta;

eudicotyledons;

core eudicots;

Rosidae;

eurosid 1;

Fabales;

Fabaceae;

Papilionoideae;

Phaseoleae;

Glycine.

1 (bases 1 to 433)

Shoenmaker, R., Keim, P.,

Vodkin, L., Erpelidg, J.,

Corvelli, V., Khanna

A., Bolla, B., Marra, M.,

Hillier, L., Kucaba, T.,

Martin, J., Beck, C.,

Wylie, T., Underwood, K.,

Stepping, B., Allen, M.,

Bowers

Y., Person, B., Swaller, T.,

Gibbons, M., Pape, D.,

Harvey, N., Schurk

R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
R., Waterston, R. and Willson, R.

# TITLE

Public Soybean EST Project  
Unpublished (1999)

## COMMENT

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 424

POLYA-No.

## FEATURES

Location/Qualifiers

1. 433

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl004-3508"

/clone\_lib="Gm-cl004"

/tissue\_type="root"

/lab\_host="XL10-Gold"

/note="Vector: pBluescript II XR; Site\_1: EcoRI; Site\_2:  
XhoI; Root cDNA. The mRNA was isolated from entire roots  
of 8 day old 'Williams' seedlings which were propagated on  
paper towels with distilled water. Stratagene's cDNA  
Synthesis Kit (catalog #200401) was used to synthesize the  
cDNA. First- strand synthesis was performed with 5-methyl  
dCTP, hence the ligated cDNA is hemimethylated.  
Stratagene's first-strand synthesis primer was used  
[GAGAGAGAGAGAGAGAGAGTAGTCGAG(T)-18]. After  
second-strand synthesis, the cDNA ends were 'polished',  
with clone pfu DNA polymerase, ligated to EcoRI adapters,  
and phosphorylated. The XhoI site within the first-strand  
synthesis primer was restricted by digestion with XhoI;  
all XhoI sites in the cDNA would be protected by their  
hemimethylated status. The cDNA constructs were  
size-fractionated with a 500bp cutoff, using GibcoBRL Life  
Technologies' cDNA Size Fractionation column. The column  
eluent was then ligated into Stratagene's pBluescript II  
XR predigested vector (pBluescript II SK(+)) that had been  
digested with EcoRI and XhoI, and phosphorylated. Both  
the white and blue colonies appear to contain recombinant  
plasmids with cDNA inserts. Blue colonies 9n-15) have been  
sequenced, and possess putative cDNA inserts. This library  
was constructed by Dr. Paul Keim & Virginia H. Coryell,  
Department of Biology, Box 5640, Northern Arizona  
University, Flagstaff, AZ 86011, Phone: 520-523-1078 (Dr.  
Paul Keim), 520-523-1372 (Virginia H. Coryell), Fax:  
520-523-7500, email: paul.keim@na.u.edu,  
virginia.coryell@na.u.edu"

## BASE COUNT

102 a 136 c 88 g 107 t

## Query Match

Best Local Similarity 22.3%; Score 433; DB 9; Length 433;

Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 GAAAGTGTAGCCACGCTCTCAACACACTTTCCTCGCGCGCGGCTCCCTCGCCACCTCC 109

|||||

DB 1 GAAAGTGTAGCCACGCTCTCAACACACTTTCCTCGCGCGCGGCTCCCTCGCCACCTCC 60

QY 110 ACCGCGGGCTCTTCAATTCGCTGAGACAACACGAGCTTCGGTGATGACTTGGCC 169

|||||

DB 61 ACCGCGGGCTCTTCAATTCGCTGAGACAACACGAGCTTCGGTGATGACTTGGCC 120

QY 170 AAGGATCTTGGTTCGACGACTCCATCAACAGCGAGCGCGCGCTCAATTCCTCCCAACAG 229

|||||

DB 121 AAGGATCTTGGTTCGACGACTCCATCAACAGCGAGCGCGCGCTCAATTCCTCCCAACAG 180

QY 230 CAAAACGAAAAACAAGACACTGATTTCTCGTCTCAAAATTCGCTACCGCTTCCTCGTC 289  
|||||

DB 181 CAAAACGAAAAACAAGACACTGATTTCTCGTCTCAAAATTCGCTACCGCTTCCTCGTC 240  
|||||

QY 290 CCGGCTCACCGCAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTTCGTCAGAGT 349  
|||||

DB 241 CCGGCTCACCGCAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTTCGTCAGAGT 300  
|||||

QY 350 CACGCGGGCTCTTCAACCTTTGTATAGTAGTCCTTGTGTGTGAATAGCGGACTCATC 409  
|||||

DB 301 CACGCGGGCTCTTCAACCTTTGTATAGTAGTCCTTGTGTGTGAATAGCGGACTCATC 360  
|||||

QY 410 ATTGAGATTTAATGAAGTATGGTTGGTTGATCAATTCGGTTTGGTTAGTTCAAG 469  
|||||

DB 361 ATTGAGATTTAATGAAGTATGGTTGGTTGATCAATTCGGTTTGGTTAGTTCAAG 420  
|||||

QY 470 TCATTGAGAGACT 482  
|||||

DB 421 TCATTGAGAGACT 433  
|||||

## RESULT 6

BI422326

LOCUS

DEFINITION

EST532992 tomato callus, TAMU Lycopersicon esculentum cDNA clone

BI422326

ACCESSION

VERSION

BI422326.1

GI:15196624

KEYWORDS

EST.

SOURCE

tomato.

ORGANISM

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

Lycopersicon.

REFERENCE

1. (bases 1 to 729)

Aucala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E.,

Li, F., Upton, J., Craven, M.B., Bowman, C.L., Ann, S., Renning,

C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.

Generation of ESTs from tomato callus tissue

Unpublished (1999)

Contact: CUGI

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>.

Location/Qualifiers

1. 729

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="cLE69N10"

/clone\_lib="tomato callus, TAMU"

/tissue\_type="callus"

/dev\_stage="25-40 days old"

/lab\_host="xLI-Blue MRF"

/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:

XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons

of seedlings 7-10 days post-germination were excised, cut

at both ends and placed on MS medium with no selection.

Mixed callus was harvested at 25 and 40 days and included

undifferentiated masses. Tomato Callus EST Library"

BASE COUNT

177 a 142 c 160 g 250 t

ORIGIN

## Query Match

Best Local Similarity 22.3%; Score 432.2; DB 13; Length 729;

Matches 533; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 785 GATACTCTGAACATGACTATCTTACACGTAAGCTTCAAGAGCTTGGCATATTTCTCG 844

|||||

DB 14 GAGAATTCGGAATCAACTACTCTTACATGTAGTTTCGAGAGTTGGCTTACTTCATG 73

Qy	845	GTGGCCCTACATATATGTACAGCCAGCTATCTCGCACACCTTATATATCGAAAGGT	904
Db	74	GTGGCTCAACATTTATGCTATACAGTTAGCTTCTCGCTCGCATTCATTCGAAAGGT	133
Qy	905	TGGTGTTCGCCAACTGTCAAGCTGATAATATTTACAGAGTTATGGGATTTATATA	964
Db	134	TGGTGGCCGCCAACTCATCAGCTGTAATTTTAGAGAGTTAATGGGATTTATCAT	193
Qy	965	GACCAATATTAATCCCATAGTACAAATTCACAGCATCTCTCAAGGGAAACCTTCT	1024
Db	194	GAGCAGTATATTAACCCGATTGTGGAAGCTCACGACATCCATTTGAAGGAAACCTTTA	253
Qy	1025	TACGCCACCCAGAGTCTCAAGCTTTCGTTCCTCAATTTATATGTCGCTCGCATG	1084
Db	254	TACGCCACCCAGAGGATTTGAAGCTTTCAGTTCCAAATTTTATATGTCGCTCGCATG	313
Qy	1085	TTCTATTGCTTTTCCACCTTTGTTAAATATCTGGCAGAGCTTCTTCGATTTGGTGAT	1144
Db	314	TTCTACAGCTCTTTCATCTTTGGTTAAATATCTTGGGGAAGTTCTGCGATTTGGGGT	373
Qy	1145	CGTGAATTTCTACAGGATTTGTTGAATGCCAAACTGTCTGAAGATTTATGAGGATGTGG	1204
Db	374	CGTGAGTTCTATAAGATTGTGTGAACGCAAAACAATTTGATGAGTATTTGAGACTTTGG	433
Qy	1205	AATATGCTGTTCACAAATGATGATCGCCACCTATATTTTCCATGTTTAAAGCAGGT	1264
Db	434	AATATGCTGTACATAAGTGTGATGCTGTCGTCACATCTATTTCCCATGCTTAAAGAAATGGC	493
Qy	1265	CTACCAAGGCTGTGCTCTTTAATGCTCTCTGCTTCTGCTTTATTCATTCATGAGCTG	1324
Db	494	ATACCTAAGGAGTTCATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	553
Qy	1325	TGCATGCTGTTCCTTGCCACATATCAAGTTGTGGGCTTTTCGGTGGAAATTTATGTTTTCAG	1384
Db	554	TGATGCTGTTCCTTGCGCTATCAAGTTTGGGCAATTCCTTGGAAATCATGTTTCAG	613
Qy	1385	GTTCCTTTGCTGTGATCACTAATATCTGCAAAATTAATTCAGAACTCAATGGTTGGA	1444
Db	614	ATTCCTTTGCTGTCACTAATCAAGCTTCTGCAAAACAAAGTTCAAAACCTCGAATGTGGGC	673
Qy	1445	AATATGATTTTGTGTCATATTCAGTATCTCTGCTCAACC	1485
Db	674	AACATGACATCTGCTGCTTTTCTGCAATTTGCTGCAACC	714
RESULT	7		
LOCUS	AV110660	1181 bp	linear
DEFINITION	Zea mays CL739_1 mRNA sequence.		
ACCESSION	AV110660		
VERSION	AV110660.1	GI:21215250	
KEYWORDS	HTC		
SOURCE	Zea mays.		
ORGANISM	Zea mays.		
REFERENCE	1	(bases 1 to 1181)	
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.		
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
REFERENCE	2	(bases 1 to 1181)	
AUTHORS	Coe,E.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
FEATURES	source		
	1..1181		
	/organism="Zea mays"		
	/db_xref="MaizeDB:632842"		
	/db_xref="taxon:4577"		

/clone="CL739_1"									
/clone_lib="Maize Mapping Project/DuPont Consensus Library"									
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"									
BASE COUNT	322 a	240 c	275 g	327 t	17 others				
ORIGIN									
Query Match	22.0%; Score 427.6; DB 11; Length 1181;								
Best Local Similarity	69.0%; Pred. No. 1.7e-58;								
Matches 587; Conservative	0; Mismatches 261; Indels 3; Gaps 1;								
QY	712	GGTGTCTTATGCACATACAACTATGATATGAGAGCCTTACCAAAATAGTTGAAAGGG	771						
Db	11	GGTCTCTTATGCACATACAAATTTATGATATAAGGGTATTGTCCAAAAGTACTGAGAAGG	70						
QY	772	AGAAGCAGCTGCTCGATCTCTGAACATGGACTATCCTTACAAGCTAAGCTTCAAGAGCTT	831						
Db	71	TGCTGCATATGGAATTTATGTCGATCTCGAATATGAAGATCCAACTTTAAAGTCT	130						
QY	832	GGCATATTTCTGTTGCCCTACATTTATGTTACCAGCCAAGCTATCTCGCACACCTTA	891						
Db	131	AGTGTACTTCACTGTGCCCAACACTTTGTTACCAGCCAACCTTATCCTCAAACTACATG	190						
QY	892	TATTCAAAGGGTGGTGTGTTTCCCAACTTGTCAAGCTGATATATTTACAGAGTTAT	951						
Db	191	TATTGAAAGGGTGGGTGACCCAGCAACTCATAAAGTGGTGGNNNNNACAGGCTTGAT	250						
QY	952	GGGATTTTAAATAGACCAATATATTAATCCATAGTACAAATTCACAGCATCTCTCAA	1011						
Db	251	GGGCTTCAATATTCAGCAATATATAACCCAAATTTGAAGAAATTCACCAACTCCACTGAA	310						
QY	1012	GGGAAACCTCTTTACGCCACCGAGAGTTCCTGAAGCTTTCTGTCTCCAAATTTATATGT	1071						
Db	311	AGGGAANNNNNGAATGCTATAGAAAGAGTCTTAAACTCTCAGTGCCAACTATATATGT	370						
QY	1072	GTGGCTCTGATGTTCTATGCTTTTTCACCTTTGGTAAATATCCTGGCAGAGCTTCT	1131						
Db	371	ATGGCTTTGCATGTTCTATTGCGNNNNNCAATTTATGGCTGAACATTTGTAGCTGAACCTCT	430						
QY	1132	TCGATTTGGTGTGATCGTGAATTCCTACAAAGATTGGTGAATGCCAAACTGTCGAAGATTA	1191						
Db	431	CTGTTTCCGGTGACCGTGAATTCCTATAGGACTGGTGAATGCCAAACTGTTGAAGAGTA	490						
QY	1192	TTGGAGGATGTGGAATATGCTGTTTCACAAATGGATGATCGGCCACCTATATTTTCCATG	1251						
Db	491	CTGGAGGATGTGGAACATGCTTGTTCATAGTGGATCATCAGACACATATATTTTCCATG	550						
QY	1252	TTTAAGGCACGGTCTACCAAGGCTGCTGCTTTTAAATTCCTTCTGCTGTTTCTGCTTT	1311						
Db	551	TATAAGGAAAGGGCTTTTCCAGGGGTGATGCTATCTTAATCTCGTTTCTGCTTTCAGCTGT	610						
QY	1312	ATTCCATGAGCTGTGATTCCTGTTTCCCTCCACATATTTCAAGTTGTGGGCTTTCGGTGG	1371						
Db	611	ATTCCATGAGATATGATTTGCGGTGCGGTGCCACATTTTCAATTTCTGGGCATTTTCTGG	670						
QY	1372	AATTATGTTTCAGGTTCTCTTTGGTCTTGATCACTAATATCTGCAAAATAAATTCAGAAA	1431						
Db	671	GATCATGTTTCAGATACCGTGGTATTCTTGACAAATATCTCCATGCTACGTTTCAAGCA	730						
QY	1432	CTCAATGGTTGAAATATGATTTTGGTTCATATTTTTCAGTATCTTGGTCAACCTATGTG	1491						
Db	731	TGTAATGGTGGGCAACATGATATTTTGGTTC---TTTCAGTATAGTCGGACAGCCGATGTG	787						
QY	1492	TGTACTGCTATACTACCATGACTTGTGAATAGGAAGGCAAACTTTGACTGAAGCTACGG	1551						
Db	788	TGTCTCTCTATCTACTACCATGACGTCATGAACAGGCGCCAGGCAAGTAGATAGTTTCGG	847						
QY	1552	CCATTACATTT	1562						

```
Db 848 CAGAGACATGT 858
|||||
RESULT 8
BQ124305
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BQ124305
EST609881 GLSD Medicago truncatula cDNA clone pGLSD-34N12, mRNA
sequence.
BQ124305
EST.
BQ124305.1 GI:20176267
barrel medic.
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 572)
Grusak M.A., Samac, D., Town, C.D., Van Aken, S., Utterback, T., Cheung
, F. and Fraser, C.M.
ESTs from late stage developing seeds of Medicago truncatula
Unpublished (2002)
Contact: Grusak, M.A.
USDA/ARS Children's Nutrition Research Center
Baylor College of Medicine
1100 Bates Street, Houston, TX 77030-2600, USA
Tel: 713 798 7044
Fax: 713 798 7078
Email: mgrusak@bcm.tmc.edu
TIGR sequence name: MTRBB78TK More information is available at:
www.medicago.org
Seq primer: Slnod (CTA gaa cta gtg gat CC).
Location/Qualifiers
1..572
/organism="Medicago truncatula"
/cultivar="A17"
/db_xref="taxon:3880"
/clone_lib="GLSD"
/tissue_type="Immature seeds"
/dev_stage="25 to 35 days after pollination"
/lab_host="XLOLR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Immature seeds, collected from pods ranging in age
from 25 to 35 days after pollination, were harvested from
greenhouse-grown plants. Seed were removed and separated
from pod walls and immediately frozen in liquid nitrogen.
Seeds throughout the age range were pooled for mRNA
extraction. cDNA was prepared from polyA+ enriched RNA.
The cDNA was directionally ligated into the Unizap XR
vector from Stratagene and packaged using Gigapack III
Gold packaging extracts. Plasmids containing cDNA inserts
were excised from the recombinant lambda-zap phage using
Ex-assist helper phage and propagated in XLOLR cells."
BASE COUNT 160 a 112 c 105 g 195 t
ORIGIN
Query Match 21.9%; Score 424.6; DB 14; Length 572;
Best Local Similarity 84.9%; Pred. No. 7.1e-58;
Matches 488; Conservative 0; Mismatches 84; Indels 3; Gaps 1;
QY 519 TTCTTTCGCGCTTTATAGTGAGAGAGTTGGCACAACGGAAGTGTATACCGGAACACG 578
|||||
Db 1 TTCTTTCGCGCGCTTTATATACAAAGTTGGCACAACGGAAGTGTATATGAACATA 60
QY 579 TTCTTTCGCTTACATATAAATCAATCACTCACTTCGCTTTCTATCAGTTTGTAGTTA 638
|||||
Db 61 TTCTTTCGCTTACATATAAATCAATCACTCACTTCGCTTTCTATCAGTTTGTAGTTA 120
QY 639 TTCTCAGGTGTATCTGCTTTGTATCAGGTGTACAGTTTATGCTGTTTCTCTGTGTTG 698
|||||
Db 121 TCCTCAGGTGTATCTGCTTTGTGTCAGGCTGCACATGATGCTACTGCTTGTCATTG 180
|||||
```

```
Qy 452 TTTTGGTTTGTAGTTCAAAGTCAATGAGACATGGCCCTTTTCATGCTGTGTCCTTCTCT 511
Db 1 TTTTGGTTTGTAGTTCATATCATGCGAGACTGGCCACTGCTAAATGTTGCTTAGTCTA 60
Qy 512 GTGGTATTTCTTTCGCGCTTTATAGTGAGAAAGTTGGCACAAGGAAAGTATACCC 571
Db 61 CCCATATTTCCCTTGGTGCAATTTGCACTCGAAAAGTTGGCATTCACAAATCTCGTTAGT 120
Qy 572 GAACCAAGTGTGTTGTACTCATATATCATCTACCTCAACTGCTGCTTTTCATCCAGTT 631
Db 121 GATCCTGTACTACCTGTTTTCACATCTCTTTTACAACTTGAATTTGATATCCAGTG 180
Qy 632 TTAGTATTTCTCAGTGATTTCTGCTTTTGTATCAGGTGTCAGCTTAATGCTGTTTTCT 691
Db 181 CTCGTGATTTCTTAAGTGATTTCTGCACTTTTATCAGGCTTTGTGTGATGTTTATGCTC 240
Qy 692 TGTGTTGTATGTTTAAATTTGGTCTTTATGCAATACATAAATATGATATGAGACACTT 751
Db 241 TGCATTTGTTGGCTGAAGCTTTGTATCTTTTGCACATACAAAACCATGATATAGAAGAACTG 300
Qy 752 ACCAAATAGTTGAAAGGGAGAGCACTGCTCGATACCTCTGAACATGCACTATCCCTTAC 811
Db 301 ATCACAAGCGCAGAGAGTTGATATGAATGAACTGACCGCGCTGGCATAGATATTTACAA 360
Qy 812 AACGTAAGCTTCAAGAGCTTGGCATATTTCTGTTGCTGCTTACATATGTTTACAGCCA 871
Db 361 GCTCCAACCTTTGGGAGTCAACATATCTATGATGCTCGACACTCTGTTATCAGCCA 420
Qy 872 AGCTATCTCCACACACTTATATTCGAAGAGTTGGTGTGTTTCCGCAACTGTTCAAGCTG 931
Db 421 AGTTATCTCCGAACACTTATGTTAGAAAGGTTGGCTGCTGCTGCAAGTTTATCTATAC 480
Qy 932 ATAATATTACAGAGTTATGGGATTTATATAGACCAATATATTAATCCCATAGTACAA 991
Db 481 TTGATATTTACTGCTCCAGGATTCATATGAGCAATACATAATCCATTTGTTG 540
Qy 992 AATTCACAGCATCTCTCAAGGAACTTCTTTACGCCACCGAGAGTTCTGAAGCTT 1051
Db 541 AACTCTCAACATCCATGATGGAGGATTAAGTATGCTGAGAGACTGTTTGAAGCTC 600
Qy 1052 TCTGTTCACAAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1111
Db 601 TCAATACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 1112 AATATCTGCGAGAGCTTCTCGATTTGGTATGCTGGAATTCACAGGATTTGTTGAAT 1171
Db 661 AACATCTTGTGAGATCTTCTGATTTGGTACCGAGAAATTCACAAAGACTGGTGGAT 720
Qy 1172 GCCAAACTGTCGAAGATTTATGGAGGATGGAATATGCTGTTCAAAATGATGATC 1231
Db 721 GCAAGACAATTTGATGAGTACTGGAGAAAATGGAACATGCTGTCATATAATGATGTT 780
Qy 1232 CGCCACCTATATTTCCATGTTTAAAGCACGCTTACCAAGGCTGCTGCTGCTTTTAAAT 1291
Db 781 CGTATATATATTTCCCTGTCATCGAAATGATATACAAAGGAAGTGTGCTGTTTATA 840
Qy 1292 GCCTCTCTGTTTCTGCTTTATTCATGAGCT 1323
Db 841 TCGTTCCTTGTGCTGCTACTTCATGAGGT 872

RESULT 10
BQ98922
LOCUS
DEFINITION BQ98922 741 bp mRNA linear EST 22-AUG-2002
OQS2023.yg ab1 OG-EFGHJ lettuce serriola Lactuca sativa cDNA clone
OQS2023, mRNA sequence.
ACCESSION BQ98922
VERSION BQ98922.1 GI:22433318
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae; Lactuca.

1 (bases 1 to 741)

Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.

Lettuce and Sunflower ESTs from the Compositae Genome Project <http://compenomics.ucdavis.edu/>

Unpublished (2002)

Contact: Alexander Kozik [R.W.Michelmore] Department of Vegetable Crops, R.W.Michelmore Lab University of California at Davis (UCD) Asmudson Hall, UCD, Davis, CA 95616, USA Tel: 1-(530)-742-1742 Fax: 1-(530)-752-9659 Email: akozik@ucdavis.edu [michelmore@ucdavis.edu] belongs to contig OG\_CA\_Contig2177, see <http://cgdb.ucdavis.edu/> for details.

Plate: OGS20 row: J column: 23.

#### FEATURES

source

Location/Qualifiers

1..741

/organism="Lactuca sativa"

/cultivar="L.serriola"

/db\_xref="taxon:4236"

/clone="OGS2023"

/clone\_lib="OG-EFGHJ lettuce serriola"

/lab\_host="E.coli"

/note="Vector: pBRCDNA5f1AB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgdb.ucdavis.edu/>

TAG LIB=OG-EFGHJ lettuce serriola

TAG TISSUE=flowers pre-fertilized

TAG\_SEQ=GGTTGACGGG"

BASE COUNT 197 a 145 c 144 g 255 t

Query Match 19.0%; Score 368.2; DB 14; Length 741;

Best Local Similarity 70.3%; Pred. No. 5 3e-49;

Matches 507; Conservative 0; Mismatches 213; Indels 1; Gaps 1;

Qy 267 AATTGCGCTACCGTCTTCCGTCACCGGCTCTTCAACCTTTGTATAGTCTCTG 326

Db 20 ATTATGCGTATCGACCGTCTTCTCCAGCTATCGAGAAATTAAGAGTCTCTCTAAGTT 79

Qy 327 CCGACACTATTTTCGTCAGAGTACGCGGCTCTTCAACCTTTGTATAGTCTCTG 386

Db 80 CTGACGCCATTTTCAAGCAGAGTATCGAGACTCTTTAAGCTTTGCATAGTCTCTAG 139

Qy 387 TTGCTGTGAATCGCGACTCATCATTTGAGAAATTAAGTATGTTGGTTGATCAAT 446

Db 140 TTGCAGTCAATGGAGACTCATCATTTGAGAAATTAAGTATGTTGGTTGATCAAT 199

Qy 447 CTGCTTTTGGTTTGTAGTCAAGTATGAGAGACTGGCCCTTTTTCATGTTGTTCTTT 506

Db 200 CCAATTTTGGTTTCAAGATTTGAGAGATTTGCGCACTCTTATGTTGCTGCTCA 259

Qy 507 CTCTGTGTTATTTCTTTCGCTGCTTTATAGTGGAGAGTTGGCACAAGGAGTGA 566

Db 260 CTCITCCAAATCTTCCCGTCTTACTGCTCTATATTTGTGAAATTAAGTGAACAAACCTA 319

Qy 567 TACCCGAACCAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 626

Db 320 TTTCAGACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 379

Qy 627 CAGTTTTAGTTATTTCTCAGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 686





/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="cLEC69M10"

/clone\_lib="tomato callus, TAMU"

/tissue\_type="callus"

/dev\_stage="25-40 days old"

/lab\_host="XLI-Blue MRF"

/note="vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2: XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons of seedlings 7-10 days post-germination were excised, cut at both ends and placed on MS medium with no selection. Mixed callus was harvested at 25 and 40 days and included undifferentiated masses. Tomato Callus EST Library"

BASE COUNT  
ORIGIN

137 a 102 c 120 g 180 t

Query Match 16.68; Score 321.8; DB 13; Length 539;

Best Local Similarity 75.8%; Pred. No. 1.3e-41;

Matches 398; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 785 GATACCTCGAATGAGCTATCCCTACAACTGAAGCTTCAAGAGCTTGGCATATTTCTCTG 844

Db 14 GAGATTCGGAATCACTACTCTTACACTGTTAGTTTCGAGAGTTGGCTTACTTCATG 73

QY 845 GTTCCCTACATATGTTACCAAGCAAGCTATCCTCCACACCTTATATTCGAAAGGTT 904

Db 74 GTTGCTCAACTTTATGCTATCAGCTTAGCTATCCTCGCTGTCATCCATTCGGAAGGTT 133

QY 905 TGGTGTTCGCCCACTGTCAAGCTGATATATTTACAGGAGTTATGGGATTTATATA 964

Db 134 TGGCTGGCCGCCCACTCATCAAGCTGGTAATTTTACAGGATTAATGGGATTTATCAT 193

QY 965 GACCAATATATTAATCCATAGTACAAATTCACAGATCTCTCAAGGGAAACCTTCTT 1024

Db 194 GAGCAGTATATTAACCGATTGCGAAGCTCAGACATCCATTGAAGGAAACCTTTTA 253

QY 1025 TACGCCACCGAGAGGTTCTGAAGCTTTCTTCCAAATTTATATGTTGCTCTGCATG 1084

Db 254 TACGCCATCGAGAGGTTATGAAGCTTTTCAGTTCCAAATTTTATATGTTGCTCTGCATG 313

QY 1085 TTCTATTGCTTTTCCACCTTTTGGTTAAATCTCTGCGAGAGCTTCTTCGATTTGGTGAT 1144

Db 314 TTCTACAGCTCTTTTCATCTTTGTTAAATATATCTTGGGAAGTTCTGCGATTTGGGGAT 373

QY 1145 CGTGAATTTCAAGAGATTTGTTGAATGCCAAACTGTGGAAGATTATTTGAGGATGTGG 1204

Db 374 CGTGAGTTCTATAAAGATTGTGGAACCCAAACAATTTGATGATATGAGACTTTGG 433

QY 1205 AATATGCTGTTCACAAATGATGATCGCCACCTATATTTTCCATGTTTAAAGCAGGTT 1264

Db 434 AATATGCTGTACATAAGTGGATGGTTCGTCCATCTATTTCCCATGCTTAAGAAATGGC 493

QY 1265 CTACCAAGAGCTGTCTCTTTTAAATGCTTCCTGCTGCTTCTGCT 1309

Db 494 ATACCTAAGGAGTGTAAGTGGTATCTCTTTCTTTATATCTGCT 538

RESULT 13

BE247899

LOCUS BE247899 582 bp mRNA linear EST 13-JUL-2000  
DEFINITION NF038D11DTR1091 Drought Medicago truncatula cDNA clone NF038D11DTR 5', mRNA sequence.

ACCESSION BE247899

VERSION BE247899.1

KEYWORDS GI:9118005

SOURCE EST.

ORGANISM barrel medic.

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

REFERENCE 1 (bases 1 to 582)

# AUTHORS

Torrez-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,

Flores, H.R., Inman, J.T., Wellner, J.W. and May, G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula drought library

Unpublished (2000)

Contact: May GD

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7391

Fax: 580 221 7380

Email: gdmay@noble.org

Medicago Genome Initiative accession: MGI:S:25730

Insert Length: 682 Std Error: 0.00

Plate: 038 row: D column: 11

Seq primer: TCACACAGGAACAGCTATGAC.

Location/Qualifiers

1. 682

/organism="Medicago truncatula"

/db\_xref="taxon:3880"

/clone="NF038D11DTR"

/clone\_lib="Drought"

/tissue\_type="Plantlets"

/dev\_stage="Pooled timepoints"

/note="Vector: Lambda Zap; Contains a mixture of entire

plantlets harvested in a series of days-post-watering

timepoints."

BASE COUNT 169 a 158 c 148 g 207 t

ORIGIN

Query Match 15.8%; Score 307.4; DB 10; Length 682;

Best Local Similarity 73.3%; Pred. No. 2.3e-39;

Matches 467; Conservative 0; Mismatches 151; Indels 19; Gaps 5;

QY 13 TCAGTCTCTCTCCATGCGGATTTCCGATGAGCCTTGAAGAGTGTAG-----CCAC 63

Db 50 TTGTTTTTATTTCTCCGATGGCGATTTCCGAGACCCCTCACGGCCGCTCCGGAACCCAAA 109

QY 64 TGCCTCTCAACACATCTTC---CCTGGCGCGCGCTCCCTCCGCCACCTCCACCGCGCGCT 120

Db 110 GGTCTCTCAACGACTCTCTGTTTACGTCGACGCGCGCGCGCTCCGTCACGGGCTT 169

QY 121 CTTCAATTCGCTGAGAC---AACCCAGCAGATTCGCGGTGATGACTTGGCCAAAGGATTC 177

Db 170 CTTGGAATCTGCTGCCACGCGTGGAAAGCGAGATTCGCGGTGAAGAGTCTGTTCCGCGTCA 229

QY 178 TGGTTCGACGACTCCATCACAGCAGCGCGCGCTCAATTTCCCAACAGCAACGCA 237

Db 230 TAGTTCGAGTACTCTATCAGCAGTGACCAAAACATCGCTGCTGCC---AATTCGATCA 286

QY 238 AAAACAAGACACTGATTTCTCCGCTCCCTCAAAATTCGCTTACCGTCTTCCGTCGCCGCTCA 297

Db 287 AATCGAGGCGCTGATATTTCTCGATCAAAATTCACATTACCGGCTTCTGTTCCGCGTCA 346

QY 298 CCGCAAGAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTTCCGTCAGAGTACGCGGG 357

Db 347 TCGTAGAAATCAAGGAAAGTCCACTCAGCTCGATAAATTTTCGCTCAGAGTCTATGCTGG 406

QY 358 CCTCTCAACCTTTGTATAGTAGTCTTGTGCTGTGAATAGCCGACTCATCTTGAAGAA 417

Db 407 CCTCTTCAACCTCTGTATAGTAGTCTTGTGCTGTGAATAGCCGACTCATCTTGAAGAA 466

QY 418 TTTAATGAAGTATGGTGGTTGATCAAAATCTGGCTTTTGGTTTATGTTTCAAGTCAATTGAG 477

Db 467 TTTAATGAAGTACGTTGGTTGATTAGACTGGCTTTTGGTTTATGTTTCAAAATCAATTGAG 526

QY 478 AGACTGGCCCTTTTCATGATGTTCTTCTTCTTGTGTTGTTTTCCTTTTCGCTGCGCTTAT 537

Db 527 AGATTGGCCCTCTTCTATGTTCTTCTTAGTCTTGTGAATATTTCTTTTGGCGGCTTTAT 586

QY 538 AGTGGGAAGTTGGCAACAGGAGTCTATACCGACAGCTGTTGTTTGTACTTTCATAT 597

Db 587 AGTGGAAAGTTGGCAACAC--GAAGTGTATTTATGAACATATTTGGTGTCTACTTTCATA 645



QY 598 AATCATTAACCTCAACTCGCTTTTCATCCAGTTTAA 634  
 Db 646 TAATCATTCGACAACTCGAGTCTCTCTATCTCTATTTA 682

## RESULT 14

BI308446

LOCUS

DEFINITION BI308446 763 bp mRNA linear EST 20-JUL-2001  
 EST529856 GPOD Medicago truncatula cDNA clone pGPOD-5N2 5' end,  
 mRNA sequence.

ACCESSION

BI308446

VERSION

BI308446.1 GI:14982773

KEYWORDS

EST.

SOURCE

barrel medic.

ORGANISM

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

REFERENCE

1 (bases 1 to 763)

AUTHORS

Grusak M.A., Samac, D.A., Town, C.D., Van Aken, S., Utterback, T., Cho,  
 J. and Fraser, C.M.

TITLE

ESTs from developing reproductive tissues of Medicago truncatula

JOURNAL

Unpublished (2001)

COMMENT

Contact: Michael A. Grusak

USDA/ARS Children's Nutrition Research Center

Baylor College of Medicine

1100 Bates Street, Houston, TX 77030-2600, USA

Tel: 713-798-7044

Fax: 713-798-7078

Email: mgrusak@bcm.tmc.edu

BI395301e

TIGR sequence name: MTOAP73TK

More information is available at: www.medicicago.org

Seq primer: SKmod (CTA GAA CTA gtg gat CC).

Location/Qualifiers

1..763

/organism="Medicago truncatula"

/cultivar="Al7"

/db\_xref="taxon:3880"

/clone="pGPOD-5N2"

/clone\_lib="GPOD"

/tissue\_type="Immature pod walls"

/dev\_stage="Immature pods, ranging in age from 15 to 30  
 days after pollination"/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
 XhoI; Immature pods, ranging in age from 15 to 30 days  
 after pollination, were collected from greenhouse-grown  
 plants. At harvest, seeds were removed from pods and  
 isolated pod walls were collected and immediately frozen  
 in liquid nitrogen. Pod walls were pooled for mRNA  
 extraction. cDNA was prepared from polyA+ enriched RNA.  
 The cDNA was directionally ligated into the Unizap XR  
 vector from Stratagene and packaged using Gigapack III  
 Gold packaging extracts. Plasmids containing cDNA inserts  
 were excised from the recombinant lambda-zap phage using  
 ex-assist helper phage and propagated in XL049 cells."

BASE COUNT

198 a 174 c 155 g 236 t

ORIGIN

source

Query Match

Best Local Similarity

Matches 444; Conservative

0; Mismatches 135; Indels 18; Gaps 4;

QY 13 TCGGTCTTCTCCAACTGGGATTTCCGATGAGCTGAAGTGTAG-----CCAC 63

Db 170 TTGCTTTTATTTCCGATGGGATTTCCGACGACCTCAGCGCGCTCGGAACCCAA 229

QY 64 TCGTCTCAACCACTCTTC-----CCTGCGCGCGCTCCCTCCGACACCTCCACCGCGCGCT 120

Db 230 GGTCTCAACGACTCTCTGTTTAGTCTGCGAGCGCGGCGGAGCGCTCTCGTCAGCGGCT 289

QY 121 CTTCAATTCCGCTGAGAC---AACCAACCAGACAGTTCGGTGATGACTTGGCCAAAGATTTC 177  
 Db 290 CTTGAATCGTCTGCCACGGTGGAAAGCGAGAGTTCGGTGAGAGTTCGCTGAAGGATTC 349  
 QY 178 TGGTTCGCGAGCTCCATCAACAGCGACGACGCGCGCTCAATTCCCAACAGCAACGA 237  
 Db 350 TAGTTCGCGATGACTCTATCAGAGTGACCAAAACATCGCTGCTGCC--AAATCTGATCA 406  
 QY 238 AAAACAAGACACTGATTTCTCCGTCCTCAAAATTCGGCTACCGCTTCGGTCCGCCCTCA 297  
 Db 407 AAATCGAGGCGCTGATATTTCTTCGATCAAAATTAATCACTTACCGGCTCTGTTCCCGCTCA 466  
 QY 298 CGCAAAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTCCCGTCAGAGTACCGGG 357  
 Db 467 TCGTAGAATCAAGAAAGTCCACTCAGCTCCGATAACATTTTCCGTCAGAGTCACTGG 526  
 QY 358 CTTCTTCAACCTTTGTATAGTAGTCTTGTGCTGTGAATAGCGGACTCATCATTCAGAA 417  
 Db 527 CTTCTTCAACCTCTGTATAGTAGTCTTGTGCTGTGAATAGCGGACTCATCATTCAGAA 586  
 QY 418 TTTAATGAAGTATGTTGTTGATCAAAATCTGGCTTTTGGTTTAGTTCAAAGTCATTTAG 477  
 Db 587 TTTAATGAAGTACGGTGTGTTGATTAAGACTGGCTTTTGGTTTAGTTCAAATCATTTAG 646  
 QY 478 AGACTGGGCGCTTTTCATGTGTTGCTTCTCTGTGGGTATTTCCCTTCGTCGCCCTTAT 537  
 Db 647 AGATTGGCGCTCTTTCATGTGTTGCTTCTCTGTGGGTATTTCCCTTCGTCGCCCTTAT 706  
 QY 538 AGTGGAGAAGTTGGGCAACAGGAAGTGTATACCGCAACAGTTGTTGTGTACTTCA 594  
 Db 707 AGTGGAAAAGTGGGCAACAGGAAGTGTATTTATGAACATATTTGGTGGTCTACTTCA 763

## RESULT 15

BJ322102/c

LOCUS

DEFINITION

BJ322102 Y. Ogiwara unpublished cDNA library, wh\_yf Triticum

aestivum cDNA clone wh\_yf23117 3', mRNA sequence.

ACCESSION

BJ322102

VERSION

BJ322102.1 GI:20125036

KEYWORDS

EST.

SOURCE

bread wheat.

ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 Triticeae; Triticum.

REFERENCE

1 (bases 1 to 729)

AUTHORS

Ogiwara, Y. and Murai, K.

TITLE

Expressed genes in Triticum aestivum

JOURNAL

Unpublished (2002)

COMMENT

Contact: Tadasi Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..729

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="wh\_yf23117"

/tissue\_type="spikelet at early flowering"

/dev\_stage="Feekes' scale 6"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
 Site\_1: EcoRI; Site\_2: XhoI; Plants were grown under  
 hydroponic conditions at UC Davis, salt stressed for 12  
 hours, and for 7 days, then dissected and frozen (Akhunov  
 in J. Dvorak Lab). Total RNA was prepared from sheath  
 tissue, equal quantities of RNA were pooled from the two  
 samples, polyA was purified from the pooled RNA, a cDNA

library was made, and the cDNA clones were in vivo excised to give pbluescript phagemids in the TJ Clouse lab at the University of California, Riverside (Akhunov, Chin, Choi, Fenton, Kianian, Otto, Simons, Zhang). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors).

BASE COUNT  
ORIGIN

237 a 146 c 153 g 193 t

Query Match 15.2% Score 295.8; DB 13; Length 729;  
Best Local Similarity 68.7%; Pred. No. 1.5e-37;  
Matches 436; Conservative 0; Mismatches 197; Indels 2; Gaps 2;

QY 914 CGCCAACCTTGTCAGCTGATAATATTTACAGGAGTTATGGGATTTATAATAG-ACCAATA 972  
Db 729 CGTCAAAATATTCTGCACTTGCATATTTACAGGTATTCAAGGTTTCATTATTGAAGCAGTA 670

QY 973 TATTAAATCCCATAGTACAAAATTC-ACAGCATCTCTCAAGGGAACCTCTTTACGCCA 1031  
Db 669 CATAAATCCCAATCGTCGTGAACATCTCCAACATCCATTGAAAGCGGACCTCTGAATGCTG 610

QY 1032 CCGAGAGAGTTCTGAAGCTTTCTGCTCCAAATTTATATGCTGTGCTCTGCAATGTTCTATT 1091  
Db 609 TAGAGACTGTTCTGAGGCTCTCATACCAATGTTTACTTAAGGCTTTCGATGTTCTATT 550

QY 1092 GCTTTTCCACCTTTGGTTAAATATCCTGGCAGAGCTTCTTCGATTTGGTGATCGTGAAT 1151  
Db 549 GCTTTTCCATCTCTGGTTAAATATCTGCTGAGATTTCTCGTTTCGGTGACCGTGAAT 490

QY 1152 TCTACAAGGATGCTGGAAATGCCAAACTGTCGAAGATTATTGGAGGATGCGGAATATGC 1211  
Db 489 TCTACAAGACTGGTGGAAATGCAAAACAATTGATGAGTATTGGCGAAATGGAACATGC 430

QY 1212 CTGTTTCAAAATGGATGATCGCCACCTATATTTTCCATGTTTAAAGGCACGCTCTACCA 1271  
Db 429 CCGTGCATAAATGGATGTTCTGTCATATATATTTTCTTGCATCGGAAGTGGTATATCGA 370

QY 1272 AGGCTGCTGCTCTTTTAAATGCTTCTGCTGCTTTTATTCATGAGCTGTGCATTTG 1331  
Db 369 AGAAGTTGCTGTTTTTGTATCATTTTGTATCTGCGGTGCTCCATGAGCTTGTGTTG 310

QY 1332 CTGTTCTCTGCCACATATTCAGTTTGGGCTTTTCGGTGGAAATATGTTTCAGGTTCTTT 1391  
Db 309 CTGTCCTCTGCCGAATCTCTGAAGTCTGGGCAATTCCTAGGGATCATGTCGAGATCCCC 250

QY 1392 TGGTCTTGATCACATAATATCTGCAAAATAAATCAGAACTCAATGTTGAAATATGA 1451  
Db 249 TTATCACATTGACATCATGCTCAAAAGCAAAATTCAGGGATACAAATGCCGCGCAACATGA 190

QY 1452 TTTTGTGGTTCATATTCAGTATTCCTTGGTCAACCTATGTGTGTACTGTCTATPACTACCATG 1511  
Db 189 TATTGTGTTCTTTTCTGCTATCTACGCCAGCCCTATGTGGGTTCTCTCTACTACTACCATG 130

QY 1512 ACTTGATGAATAGCAAGGCAAACTTGACTGAAC 1546  
Db 129 ATGTGATGAACAGGATTGGGAAGACAGAAATAGAC 95

Search completed: February 21, 2003, 18:31:26  
Job time : 2564 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	719.8	37.1	1942	4	US-09-326-203A-1	Sequence 1, Appli
2	229.8	11.8	629	4	US-09-103-754A-3	Sequence 3, Appli
3	199.4	10.3	234	4	US-09-326-203A-3	Sequence 3, Appli
4	194.6	10.0	275	4	US-09-326-203A-5	Sequence 5, Appli
5	187.2	9.6	267	4	US-09-326-203A-4	Sequence 4, Appli
6	170.2	8.8	1650	4	US-09-103-754A-2	Sequence 2, Appli
7	162.6	8.4	1976	3	US-09-165-042-2	Sequence 2, Appli
8	159.8	8.2	1766	4	US-09-326-203A-15	Sequence 15, Appli
9	159.8	8.2	1766	4	US-09-326-203A-16	Sequence 16, Appli
10	159.4	8.2	1895	4	US-09-326-203A-14	Sequence 14, Appli
11	124.4	6.4	253	4	US-09-326-203A-7	Sequence 7, Appli
12	114.2	5.9	254	4	US-09-326-203A-8	Sequence 8, Appli
13	108.2	5.6	262	4	US-09-326-203A-9	Sequence 9, Appli
14	92	4.7	325	4	US-09-326-203A-10	Sequence 10, Appli
15	65.6	3.4	4011	1	US-08-121-057-3	Sequence 3, Appli
16	65.6	3.4	4011	2	US-08-509-187D-3	Sequence 3, Appli
17	65.6	3.4	4011	2	US-09-121-396-3	Sequence 3, Appli
18	65.6	3.4	4011	5	PTC-US93-09704A-3	Sequence 3, Appli
19	65.6	3.4	4079	1	US-08-121-057-2	Sequence 2, Appli
20	65.6	3.4	4079	2	US-08-509-187D-2	Sequence 2, Appli
21	65.6	3.4	4079	2	US-09-121-396-2	Sequence 2, Appli
22	65.6	3.4	4079	5	PTC-US93-09704A-2	Sequence 2, Appli
23	60.4	3.1	3238	4	US-08-123-934A-5	Sequence 5, Appli
24	60.4	3.1	3238	5	PTC-US94-10080-5	Sequence 5, Appli
25	60	3.1	1582	3	US-08-545-196B-10	Sequence 10, Appli
26	60	3.1	1582	3	US-08-545-196B-12	Sequence 12, Appli
27	59.2	3.0	2025	4	US-09-149-476-316	Sequence 316, Appli

Db 821 TGAGAAATGGTACTTCAGAAATACATATACAGAACCTGTGTGTCATCTTTCTTCATATAT 880  
QY 601 CATTACCTCAACCTTCGCTTTCTATCCAGTTTTAGTATTTCTCAGGTGTGATTCGCTTT 660  
Db 881 TATCACCATGACAGAGTTTGTATCCAGTTTACGTACCCCTAAGGTGTGATTCGCTTT 940  
QY 661 TGTATCAGGTGTGACGTTTATAGTCTTTCTTGTGTGTGATGTTTAAATTTAGTGTGCTTA 720  
Db 941 TTTATCAGGTGTGACCTTTGATGCTCTCTCACTTGCATTTGTGCTAAAGTTGTTCTTA 1000  
QY 721 TGCACATACAACTATGATATGATGACCTTACCAATTTAGTTGAAAGGGGAGCACT 780  
Db 1001 TGTCTACTAGTATGACATAAGATCCCTAGCCATGACGTGATAAGG----- 1050  
QY 781 GCTCGATCTCTGAACATGACTATCTTACACAGCTTCAAGAGCTTCAAGAGCTTGGCATATTT 840  
Db 1051 -----CCAATCCTGAAGTCTCTTACTAGTTAGCTTGAAGAGCTTGGCATATTT 1099  
QY 841 CTTGTTGGCCCTTACATTTATGTTTACAGCAAGCTATCTTCGACACCTTATATTCGAA 900  
Db 1100 CATGGTGGCTCCACATTTGTTTATCAGCAAGTTATCCAGCTTCTGCATGTATACGAA 1159  
QY 901 GGCTTGGTGTTCGCAACTTCTCAAGCTGATAATTTTACAGGAGTTATGGGATTTAT 960  
Db 1160 GGCTTGGTGGCTCGTCAATTTTGCAAAATGCTGTCATATTCACCGGATTCATGGGATTTAT 1219  
QY 961 AATAGACCAATATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGGAACCT 1020  
Db 1220 AATAGAACATATATATATCTTATCTCAGGAATCTCAAGCATCTTGAAGGCGATCT 1279  
QY 1021 TCTTTACGCCACCGAGAGTCTGAAGCTTTCTGTTTCCAAATTTATATGTGTGCTCTG 1080  
Db 1280 TCTATATGCTATTAAGAGAGTGTGAAGCTTTTCAAGTCTTCAAAATTTATATGTGTGCTCTG 1339  
QY 1081 CATGTTCTATGCTTTTTCACCTTTGTTAAATATCTTCCGAGAGCTTCTTCGATTTG 1140  
Db 1340 CATGTTCTACTGCTTCTTCCACCTTTGGTTAAACATATTGGCAGAGCTTCTCTGCTTCG 1399  
QY 1141 TGATCGTGAATTTACAAGGATGTGGAAATGCCAAAGCTGCGAAGATTTATGGAGGAT 1200  
Db 1400 GGATCGTGAATTTACAAGATTTGGTGAATGCCAAAGTGTGGAGATTTACTGGAGAT 1459  
QY 1201 GTGGAATATGCTTTTCAAAATGGAATGATCGGCCACCTATATTTTCCATGTTTAAAGCA 1260  
Db 1460 GTGGAATATGCTTTTCAAAATGGAATGATCGGCCACCTATATTTTCCGCTTTCGCGAG 1519  
QY 1261 CGGTCTACCAAGGCTGCTCTCTTTTAAATGCTTCTGCTTCTGCTTTATTCATGA 1320  
Db 1520 CAAGATACCAAGACACTCGCCATTTATCATTTGCTTTTCTAGTCTCTGCACTCTTTTATGA 1579  
QY 1321 GCTGTGATTTGCTTCTTCCATATATTCAGATTTGTGGGCTTTTCGGTGGAAATATGTT 1380  
Db 1580 GCTATGATCGCATTTCTTGTGCTCTCTTCAAGCTATGGCTTTTCTTGGGATATGTT 1639  
QY 1381 TCAGGTTCTTGGCTTGTATCTAATTTATCTGCAAAATTAATTCAGAAACCTCAATGTT 1440  
Db 1640 TCAGGTGCTTGTGCTTCTATCACAACATCTACAGAAAGGTTTGG---CTCAACGCT 1696  
QY 1441 TGAATATGATTTTTTGTGCTATATTCATATTCAGTATCTTGTGCTGCTTGTGCTTGT 1500  
Db 1697 GGGAAATGATCTTCTGCTGCTATCTTCTGCAATTTTCGGAACACCGATGTGTGCTTCT 1756  
QY 1501 ATACTACCATGCTTGTGATGAATAGGAAGGCAAACTTTGATCTGAAGTACGG 1551  
Db 1757 TTATTACCGACCTGATGACGGAAGGATCGATGTCATGAAACAACTG 1807

RESULT 2

US-09-103-754A-3

; Sequence 3, Application US/09103754A

; Patent No. 6344548

; GENERAL INFORMATION:

; APPLICANT: Farese, Robert

APPLICANT: Cases, Sylvaine  
APPLICANT: Smith, Steven  
APPLICANT: Erickson, Sandra  
TITLE OF INVENTION: Diacylglycerol O-acyltran  
TITLE OF INVENTION: sferase  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bozicevic & Reed  
STREET: 285 Hamilton Avenue, Suite 200  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/103,754A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Field, Bret E  
REGISTRATION NUMBER: 37,620  
REFERENCE/DOCKET NUMBER: 6510-105p  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650 327 3400  
TELEFAX: 650 327 3231  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 629 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-103-754A-3

Query Match 11.8%; Score 229.8; DB 4; Length 629;  
Best Local Similarity 79.6%; Pred. No. 6.2e-46;  
Matches 284; Conservative 0; Mismatches 72; Indels 1; Gaps 1;

QY 888 CTTATATTCGAAAGGGTGTGGTTGTCGCGCAACTTGTCAAGCTGATATATTACAGGAG 947

Db 3 CATGTATACGGAAGGGTGTGGTGGCTCGTCAATTTGCAAACTGGTCAATATTCACCGGAT 62

QY 948 TTATGGGATTTTAAATAGACCAATATATTAAATCCCATAGTACAAAATTCACAGCATCTC 1007

Db 63 TCATGGGATTTTAAATAGACCAATATATTAAATCCATATTGTCCAGGAATCAAGCATCTCT 122

QY 1008 TCAAGGAAACCTTCTTTTACGCCACCGAGAGTTCCTGAAGCTTTTCTGTTTCCAAATTTAT 1067

Db 123 TGAAGGCGATCTTCTATATGCTATTGAAAGAGTGTGGAAGCTTTTCAGTTCCAAATTTAT 182

QY 1068 ATGTGTGGCTCTGCATGTTCTATTGCTTTTCCACCCTTTGGTTAAATATCTCGGAGAGC 1127

Db 183 ATGTGTGGCTCTGCATGTTCTATTGCTTTTCCACCCTTTGGTTAAATATTTGGCAGAGC 242

QY 1128 TTCTTCGATTTTGTGATCGTGAATTTCTACAAGGATTTGGTGAATGCCAAACTGTGCAAG 1187

Db 243 TTCTCTGCTTCGGGATCGTGAATTTCTACAAGATTTGGTGAATGCCAAAGTGTGGGAG 302

QY 1188 ATTA-TTCGAGGATGTGGAATATGCTTGTTCACAAATGGATGATCGGCCACCTATAT 1243

Db 303 ATTACTGGGAGAATGTGGAATATGCTTGTTCCTCATAAATGGATGGGTGGTCCGACATATAT 359

RESULT 3

US-09-326-203A-3

; Sequence 3, Application US/09326203A  
; Patent No. 6444876  
; GENERAL INFORMATION:

; APPLICANT: Lassner, Mike  
; APPLICANT: Ruzinsky, Diane  
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
; TITLE OF INVENTION: Acid Sequences

; FILE REFERENCE: 17045/00/WO  
; CURRENT APPLICATION NUMBER: US/09/326,203A

; CURRENT FILING DATE: 1999-06-04

; PRIOR APPLICATION NUMBER: 60/088,143

; PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/108,389

; PRIOR FILING DATE: 1998-11-12

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 234

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (24)..(60)

; OTHER INFORMATION: n at positions 24, 41, and 60 is unknown

US-09-326-203A-3

Query Match 10.3%; Score 199.4; DB 4; Length 234;

Best Local Similarity 95.7%; Pred. No. 7.8e-39;

Matches 225; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 815 GTAAGCTTCAAGAGCTTGCAATATTCCTGGTTGCCCTA-CATTATGTACAGCCCAAG 873

DB 1 GTAAGCTTCAAGAGCTTAGCATANTCTCGTTGCCCTTANCAATTATGTACAGCCCAAN 60

QY 874 CTATCTCTGCACACCTTATATCGAAAGGTTGGTTGTTTCCCAACTGTGCAAGCTGAT 933

DB 61 CTATCTCTGCACACCTTATATCGAAAGGTTGGTTTCCCAACTGTGCAAGCTGAT 119

QY 934 ATATTTACAGGAGTTATGGGATTTATATAGACCAATATATTAATCCCATGTACAAA 993

DB 120 AATATTTACAGGAGTTATGGGATTTATATAGACCAATATTAATCCCATGTACAAA 179

QY 994 TTCACAGCATCTCTCAAGGGAACTTCTTTACGCCACCGAGAGTTCTGAAG 1048

DB 180 TTCACAGCATCTCTCAAGGGAACTTCTTTACGCCACCGAGAGTTCTGAAG 234

RESULT 4

US-09-326-203A-5

; Sequence 5, Application US/09326203A

; Patent No. 6444876

; GENERAL INFORMATION:

; APPLICANT: Lassner, Mike

; APPLICANT: Ruzinsky, Diane

; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic

; TITLE OF INVENTION: Acid Sequences

; FILE REFERENCE: 17045/00/WO

; CURRENT APPLICATION NUMBER: US/09/326,203A

; CURRENT FILING DATE: 1999-06-04

; PRIOR APPLICATION NUMBER: 60/088,143

; PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/108,389

; PRIOR FILING DATE: 1998-11-12

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 275

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (192)..(263)

; OTHER INFORMATION: n at positions 192, 202, 204, 211, 222, 234, 238,

; OTHER INFORMATION: 239, 244, 245, 247, 251, 262, and 263 is unknown  
US-09-326-203A-5

Query Match 10.0%; Score 194.6; DB 4; Length 275;

Best Local Similarity 82.2%; Pred. No. 1.2e-37;

Matches 226; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

QY 1165 GTGGAATGCCAAACTGTCGAAGATTATTGGAGGATGTGGAATATGCTGTTCACAAATG 1224

DB 1 GTGGAATGCCAAACTGTTGAAGATTATTGGAGGATGTGGAATATGCTGTTCACAAATG 60

QY 1225 GATGATCCGCCACCTATATTTCCATGTTTAAGGCACGCTCTACCAAGGCTGCTCTCT 1284

DB 61 GATGATCCGCCACCTATATTTCCATGTTTAAGGCACGCTCTACCAAGGCTGCTCTCT 120

QY 1285 TTTAATTGCTTCTCCTGGTTTCTGCTTTATTCATGAGCTGTGCTATGCTGTTCCTTGCCA 1344

DB 121 TTTAATTGCTTCTCCTGG-TTCTGCTTTATTCATGAGCTGTGCTATGCTGTTCCTTGCCC 179

QY 1345 CATATTCAAGTTGGGCTTTCGGTGGGAATTATGTTTCAGGTTCCCTTGTGATAC 1404

DB 180 ACATATTCAAGTNGTGGTTTCNGNGGAATTNAGTTTCAGTNCCTTGGGTTTCNACCNN 239

QY 1405 TAATTATCTGCAAAATAAATTTCAGAACTCAATGG 1439

DB 240 AATTNNGGCNAAAAAATTCNNGAACCCCGGG 274

RESULT 5

US-09-326-203A-4

; Sequence 4, Application US/09326203A

; Patent No. 6444876

; GENERAL INFORMATION:

; APPLICANT: Lassner, Mike

; APPLICANT: Ruzinsky, Diane

; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic

; TITLE OF INVENTION: Acid Sequences

; FILE REFERENCE: 17045/00/WO

; CURRENT APPLICATION NUMBER: US/09/326,203A

; CURRENT FILING DATE: 1999-06-04

; PRIOR APPLICATION NUMBER: 60/088,143

; PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/108,389

; PRIOR FILING DATE: 1998-11-12

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 267

; TYPE: DNA

; ORGANISM: Glycine max

US-09-326-203A-4

Query Match 9.6%; Score 187.2; DB 4; Length 267;

Best Local Similarity 83.6%; Pred. No. 6.8e-36;

Matches 224; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

QY 654 CTGCTTTTGTATCAGGTGTCAGGTTAAATGCTGTTTCTTGTGTGTATGTTTAAATTTGG 713

DB 1 CTGCTTTTGTATCAGGTGTCAGGTTGATGCTATTAACTTGCATTGTGTGTTAAATTTGG 60

QY 714 TGTCTTATGCACATACAACTATGATATGAGACGCTTACCAAAATAGTTGAAAGGGAG 773

DB 61 TGTCTATATGCACATACAACTATGATATGAGACGCTTACTGTGTTGAATGAAAGGGAG 120

QY 774 AAGCACTGCTGATCTCTGAACATGGACTATCCTTACACGCTTAAGCTTCAAGAGCTTGG 833

DB 121 AAACATTACCCAATCTTTG-ATATGGAGTATCCGTACACTGTGACCTTCAGAGTTGG 179

QY 834 CATATTTCTGTTGCCCTTACATTATGTTACGCAAGCTATCCTCCGACACCTTATA 893

DB 180 CATACTTCATGTTGCTCTCATATGATATGATATGATATGATATGATATGATATGATATGAT 239

QY 894 TTCGAAAGGGTGGTTGTTTCGCCAACT 921





Db 1100 CCACTCATGTCTCAATGCTGTGGCAGAGCTCCTGCAGTTTGGAGACCGCGAGTTCTACAG 1159  
QY 1159 GGATTGTGGAAATGCCAAACCTGCGAAGATTAATTGGAGGAGTGTGAATATGCTGTGTTCA 1218  
Db 1160 GGACTGTGGAAATGCTGAGTCTGTCACTACTTTGGCAGAACTGGAAATATCCCGGTGCA 1219  
QY 1219 CAAATGGATGATCGCCACCTATATTTTCCATGTTTAAAGGCACGCTGTACCAAGGCTGC 1278  
Db 1220 CAACTGTGTCATCAGACACTTCTCAAGCCATATGCTCAGACTGGGCACCAAAATGGAT 1279  
QY 1279 TGCCTCTTTAATTCCTTCCCTGCTTCTGCTTTTATTCATGAGCTGTGCAATGCTGTTC 1338  
Db 1280 GGCCAGGACTGGGTCTTTTGGCGTCAAGCTTCTTCCATGAGTACCTAGTGAACATTC 1339  
QY 1339 TTGCCACATATCAAGTTGCTGGCTTTCCGTTGGAATATGTTTCCAGTTCCTTTGGTCTT 1398  
Db 1340 CCTGAGATGTTCCGCTCTGGGCAATTCACAGCCATGATGCTCAGGTCCACCTGGCCTG 1399  
QY 1399 GAT 1401  
Db 1400 GAT 1402

RESULT 10  
US-09-326-203A-14  
; Sequence 14, Application US/09326203A  
; Patent No. 6444876  
; GENERAL INFORMATION:  
; APPLICANT: Lassner, Mike  
; APPLICANT: Ruzinsky, Diane  
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
; TITLE OF INVENTION: Acid Sequences  
; FILE REFERENCE: 17045/00/WO  
; CURRENT APPLICATION NUMBER: US/09/326,203A  
; PRIOR FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 60/088,143  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/108,389  
; PRIOR FILING DATE: 1998-11-12  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 1895  
; TYPE: DNA  
; ORGANISM: Human  
; NAME/KEY: misc\_feature  
; LOCATION: (209)  
; OTHER INFORMATION: n at position 209 is unknown  
US-09-326-203A-14

Query Match 8.2%; Score 159.4; DB 4; Length 1895;  
Best Local Similarity 54.88; Pred. No. 6.3e-29;  
Matches 338; Conservative 0; Mismatches 276; Indels 3; Gaps 1;  
QY 789 CTCTGAACATGGACTATCCTTACAACTTAAGCTTCAAGAGCTTGGCATATTTCTGTGTTG 848  
Db 658 CGCACACCGTGAGCTACCCGGAATCTGACCTACCCGATCTCTACTTCTCTCTCG 717  
QY 849 CCGCTACATTAATGTTACAGCAGCTATCTCTGCACACCTTATATTTGAAAGGTTGGT 908  
Db 718 CCCCCACCTTGTGTACGAGCTCAACTTCCCGCTCTCCCGCATCCGGAAGCGCTTC 777  
QY 909 TGTTCGCCAACTTGTCAAGCTGATATATTTACAGGAGTATGGGATTTATATACACC 968  
Db 778 TGCTGCAGCGATCTTGAGATGCTGTCTTACCCAGCTCCAGGTGGGCGTGATCAGC 837  
QY 969 AATATAATTAATCCAGTACAAATTCACAGCATCTCTCAAGGGAA---ACCTTCTTT 1025  
Db 838 AGTGGATGGTCCCAACCATCCAGAACTCCATGAAGCCCTTCAAGGACATGGACTACTCAC 897  
QY 1026 ACGCCACCGAGAGTCTCGAAGCTTCTGTTCCAAATTTATATGCTGCTGCGATG 1085

Db 898 GCATCATCGAGCGCTCCTGGAAGCTGGCGTCCCAATACCTCATCTGGCTCATCTTCT 957  
QY 1086 TCTATTCTCTTTTCCACCTTTGGTTAAATATCCCTGGCAGAGCTTCTTCGATTTGGTGATC 1145  
Db 958 TCTACTGGCTCTTCCACTCCTGCTGAATGCCGTGGGTGAGCTCATGAGTTTGGAGACC 1017  
QY 1146 GTGAATTTCTACAAGGATTTGGGAATGCCAAACTGTGCAAGATTTATTTGGAGGATGTGA 1205  
Db 1018 GGGAGTTCTACCGGACTGTGGAACTCCGAGTCTGTCACTACTTCTGGCAGAACTTGA 1077  
QY 1206 ATATGCTGTTTCAAAATGGAATGATGCCGCCACCTATATTTTCCATGTTTAAAGGCACGTC 1265  
Db 1078 ACATCCCTGTGCACAAAGTGTGTCATCAGACACTTCTACAAGCCCATGCTTCGACGGGCA 1137  
QY 1266 TACCAAGGCTGCTGCTCTTTTAAATGCTTCCCTGCTTCTGCTTTTATTTCCATGAGCTGT 1325  
Db 1138 GCAGCAAGTGGATGCCAGGACAGGGGTGTTCCTGGCTCGGCTTCTTCCACGAGTACC 1197  
QY 1326 GCATTGCTTCTTCCACATATTTCAAGTTGTGGGCTTTTCGGTGAATTTATGTTTCAGG 1385  
Db 1198 TGGTACGCTCCCTCTCGGAATGTTCCGCTCTGGGCTTCACGGGCTCATGATGCTCAGA 1257  
QY 1386 TTCCTTTGGCTTGTATC 1402  
Db 1258 TCCACTGGCTGTGTTTC 1274

RESULT 11  
US-09-326-203A-7  
; Sequence 7, Application US/09326203A  
; Patent No. 6444876  
; GENERAL INFORMATION:  
; APPLICANT: Lassner, Mike  
; APPLICANT: Ruzinsky, Diane  
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic  
; TITLE OF INVENTION: Acid Sequences  
; FILE REFERENCE: 17045/00/WO  
; CURRENT APPLICATION NUMBER: US/09/326,203A  
; PRIOR FILING DATE: 1999-06-04  
; PRIOR APPLICATION NUMBER: 60/088,143  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/108,389  
; PRIOR FILING DATE: 1998-11-12  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 253  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-326-203A-7

Query Match 6.4%; Score 124.4; DB 4; Length 253;  
Best Local Similarity 73.2%; Pred. No. 6.8e-21;  
Matches 186; Conservative 0; Mismatches 56; Indels 2; Gaps 2;  
QY 1196 AGGATGTGGAATATGCTTGTTCACAAATGGATGATCGCCACCTATATTTTCCATGTTTA 1255  
Db 1 AGAAATGGAAATGCTGTGCATAAATGGATGTTCTGCTATATATTTTCTTGCATG 60  
QY 1256 AGGCAGGCTTACCAAGGCTGCTCTTTTAAATTCCTTTCCTGCTTCTGCTTTATTC 1315  
Db 61 CGAAATGGTATATCAAGGAAGTTGCTGTTTATATATCGTT-CTTGTTCTGCTGTTACTT 119  
QY 1316 CATGAGCTGTGCATGCTGTTTCTTCCACATATTCAGTTTGGGCTTTCGGT-GGAAT 1374  
Db 120 CATGAGTTATGTTGCTGTTCTTCCCTGCCACATCTCAAGTTCTGGGCTTTTTTAGGAAT 179  
QY 1375 TATGTTTTCAGTCTTCTTGGTCTTGTGATCACTTAATTCCTCAAAATAAATTCAGAAATC 1434  
Db 180 CATGCTTCAGATTCCTCCCTCATCATATTTGACATCATACCTCAAAATAAATTCAGTGAC 239  
QY 1435 AATGGTTGGAATA 1448



```

Db 240 AATGGTTGGCAATA 253
|||||
; TYPE: DNA
; ORGANISM: Zea mays
US-09-326-203A-9

Query Match 5.6%; Score 108.2; DB 4; Length 262;
Best Local Similarity 64.7%; Pred. No. 5.1e-17;
Matches 161; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

RESULT 12
US-09-326-203A-8
; Sequence 8, Application US/09326203A
; Patent No. 6444876
; GENERAL INFORMATION:
; APPLICANT: Lassner, Mike
; APPLICANT: Ruzinsky, Diane
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic
; FILE OF INVENTION: Acid Sequences
; FILE REFERENCE: 17045/00/WO
; CURRENT APPLICATION NUMBER: US/09/326, 203A
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,143
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/108,389
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 254
; TYPE: DNA
; ORGANISM: Zea mays
US-09-326-203A-8

Query Match 5.9%; Score 114.2; DB 4; Length 254;
Best Local Similarity 66.4%; Pred. No. 1.9e-18;
Matches 164; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 423 TGAAGTATGGTTGGTTGATCAAAATCGGCTTTTGGTTTAAAGTCAAGTCAATTCGAGAGACT 482
|||||
Db 1 TGAAGTATGGCTTATTAATAAGATCTGGCTTTTGGTTTAAAGTCAATTCGAGAGACT 60
|||||
QY 483 GCCCTTTTTCATGTGCTCTCTTCTTGTGGTATTTTCTTCCCTTTCGCTGCTTATAGTGG 542
|||||
Db 61 GCCCACTGCTAATGTGCTTCCCTTAGTCTACCCATATTTCCCTTGGTGATTTGCGAGTCG 120
|||||
QY 543 AGAAGTTGGCACAACGGAAGTGTATACCCGAAACAGTGTGTTGTGTTACTTCAATATAATCA 602
|||||
Db 121 AAAAGTTGGCAATTCACAATCTCATAGTATCCCTGCTACTACCTGTTTTCACATCCCTTT 180
|||||
QY 603 TTACCTCAACTTCGGCTTTTCTATCCAGTTTATAGTATTTCTCAGGTGTGATTCGCTTTTG 662
|||||
Db 181 TTACAACATTTGAAATTTGTATATCCAGTCTGCTGATTTCTTAAGTGTGATTTCTGCACTTT 240
|||||
QY 663 TATCAGG 669
|||||
Db 241 TATCAGG 247
|||||

RESULT 13
US-09-326-203A-9
; Sequence 9, Application US/09326203A
; Patent No. 6444876
; GENERAL INFORMATION:
; APPLICANT: Lassner, Mike
; APPLICANT: Ruzinsky, Diane
; TITLE OF INVENTION: Acyl-CoA:Cholesterol Acyltransferase Related Nucleic
; FILE OF INVENTION: Acid Sequences
; FILE REFERENCE: 17045/00/WO
; CURRENT APPLICATION NUMBER: US/09/326, 203A
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,143
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/108,389
; PRIOR FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 262

; TYPE: DNA
; ORGANISM: Zea mays
US-09-326-203A-9

Query Match 4.7%; Score 92; DB 4; Length 325;
Best Local Similarity 61.8%; Pred. No. 4.1e-13;
Matches 162; Conservative 0; Mismatches 95; Indels 5; Gaps 2;

QY 1301 GTTTCGCTTTTATTCATCAGCTGTCATTCGTTCCCTTCCCTGACACATATTCAGTTGTGG 1360
|||||
Db 19 GGTTCAGCTGTATNCCATGAGATATGTAATGCGGTGCGCTGACATATTCANATCTNGG 78
|||||
QY 1361 GCTTTTCGGTGAATTTATGTTTCAGGTTCTTTTGGTCTTCTCAGTCACTAATTTATGCAAAAAT 1420
|||||
Db 79 CATNCCNGGG--ATCATNGTTCAGATACCGNTGGNATTTTTCACAAGATATCTCCATGCT 136
|||||
QY 1421 AAATTCAGAACTCAATGTTGGAAATATGATTTTTTGGTTTCATATTCAGTATCTCTTGGT 1480
|||||
```

Db 137 ACGTTCAAGCATGTAATGGTGGCAACATGATANTTTGGNTC---TNCAGTATAGTCGGA 193  
Qy 1481 CAACCTATGTGTACTACTACTACTACCTGACTGATGATAGGAAGGCAAACTTGAC 1540  
Db 194 CAGCGATGTNNNNNATCTATATACCATGACGTGATGAAACAGCAGCGCCAGGCAAGT 253  
Qy 1541 TGAAGCTACGCCCATATACATTT 1562  
Db 254 AGATAGTNCGGCAGACATGT 275

## RESULT 15

US-08-121-057-3  
; Sequence 3, Application US/08121057  
; Patent No. 5484727  
; GENERAL INFORMATION:  
; APPLICANT: CHANG, TA-YUAN  
; APPLICANT: CHANG, CATHERINE C. Y.  
; TITLE OF INVENTION: ACYL COENZYME A: CHOLESTEROL  
; TITLE OF INVENTION: ACYLTRANSFERASE (ACAT)  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET, SUITE 510  
; CITY: BOSTON  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/121,057  
; FILING DATE:  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LAMPORT HAMMITTE, ANN.  
; REGISTRATION NUMBER: 34,858  
; REFERENCE/DOCKET NUMBER: DCI-033CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-2700  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4011 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1397..3046  
; US-08-121-057-3

Query Match 3.4%; Score 65.6; DB 1; Length 4011;  
Best Local Similarity 53.6%; Pred. No. 2.2e-06;  
Matches 162; Conservative 0; Mismatches 134; Indels 6; Gaps 1;  
Qy 1085 TTCTATTGCTTTTCCACCTTTGGTTAAATATCCTGGCAGAGCTTCTTCGATTTGGTGAT 1144  
Db 2537 TTTTTCCTTTTGGCTTGTGCTCAATGCGCTTTGCTGAGATGTTACGCTTTGGTGAC 2596  
Qy 1145 CGTCAATCTACAGGATGCTGGAATGCCAAACTGTCGAAGATTATTGGAGGATGCG 1204  
Db 2597 AGGATGTTCTATAAGGATTTGGTGAATCCACGTCATCTCCAACTATATAGAACCTGG 2656  
Qy 1205 AATATGCTGTTCACAAATGGATGATCCGC-----CACCTATATTTTCCATGTTTAAGG 1258  
Db 2657 AATGCTGTGTCATGCTGCTATATATCTATGCTTTACAAGGACTTTCCTGCTTTTC 2716

Qy 1259 CACGGTCTACCAAAAGGCTGCTGCTCTTTTAAATGCTTCCCTGGTTCCTGCTTTTATTCAT 1318  
Db 2717 TCCAAGAGATTCAAATCTGCTGCCATGTTAGCTGCTTTGCTGCTATCTGCTGTAGTACAC 2776  
Qy 1319 GAGCTGTGCATTGCTGCTTCCCTTGGCCACATATTCAGTTGTGGGCTTTTCGGTGGAAATTATG 1378  
Db 2777 GAATATGCTTTGGCTGTTTGGCTTGGCTTTTCTATCCCGTCTGCTGCTGCTCTCTTCATG 2836  
Qy 1379 TT 1380  
Db 2837 TT 2838

Search completed: February 21, 2003, 17:48:45  
Job time : 113 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 15:48:04 ; Search time 5015 Seconds  
(without alignments)  
11269.721 Million cell updates/sec

Title: US-09-856-018b-15  
Perfect score: 1942  
Sequence: 1 tagaacaacactcgtgtttt.....aaaaaaaaaaaaaaaaaaaaa 1942

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.nam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	810.2	41.7	1964	8	AF298815	AF298815 Perilla f
2	754	38.8	2090	8	AY084052	AY084052 Tropaeolu
3	744.6	38.3	2099	8	AF129003	AF129003 Nicotiana
4	719.8	37.1	1904	8	ATH238008	ATH238008 Arabidops
5	719.8	37.1	1942	6	AX090345	AX090345 Arabidops
6	719.8	37.1	1942	6	AX090349	AX090349 Sequence
7	719.8	37.1	1988	8	ATH131831	ATH131831 Arabidops
8	719.8	37.1	2005	8	AY054480	AY054480 Arabidops
9	715	36.8	1845	8	AF051849	AF051849 Arabidops
10	701.2	36.1	1537	8	AF251794	AF251794 Brassica
11	700.2	36.1	1512	8	AF164434	AF164434 Brassica
12	605.6	31.2	1446	8	AF155224	AF155224 Brassica
13	199.4	10.3	234	6	AX090350	AX090350 Sequence
14	194.6	10.0	275	6	AX090352	AX090352 Sequence
15	187.2	9.6	267	6	AX090351	AX090351 Sequence
16	170.2	8.8	1650	10	AF078752	AF078752 Mus muscu
17	170.2	8.8	1776	10	BC003717	BC003717 Mus muscu
18	162.6	8.4	1792	9	AF236018	AF236018 Cercopith
19	162.6	8.4	1976	9	AF059202	AF059202 Homo sapi
20	161.4	8.3	34948	3	CBRG33E23	AC084551 Caenorhab
21	159.8	8.2	728	10	RNO345014	AJ345014 Rattus no
22	159.8	8.2	1751	10	AF296131	AF296131 Rattus no
23	159.8	8.2	1766	6	AX090340	AX090340 Sequence
24	159.4	8.2	1493	9	BC006263	BC006263 Homo sapi
25	159.4	8.2	1895	6	AX090339	AX090339 Sequence
26	159.4	8.2	1998	9	BC015762	BC015762 Homo sapi
27	153	7.9	2074	3	AY051835	AY051835 Drosophil
28	152.2	7.8	1935	4	AY093657	AY093657 Sus scrof
29	151.4	7.8	1497	3	AF221132	AF221132 Caenorhab
30	151.4	7.8	2117	3	AF468650	AF468650 Drosophil
31	151.4	7.8	2342	3	AF468649	AF468649 Drosophil
32	151.4	7.8	11966	3	CEH19N07	292835 Caenorhabd
33	137.6	7.1	253305	3	PFMAL3P7	AL034559 Plasmodiu
34	124.4	6.4	253	6	AX090354	AX090354 Sequence
35	114.2	5.9	254	6	AX090355	AX090355 Sequence
36	108.4	5.6	645	11	PMIC2B	AL684529 Penicilli
37	108.2	5.6	262	6	AX090356	AX090356 Sequence
38	101	5.2	92822	8	AC005917	AC005917 Arabidops
39	101	5.2	97495	-8	AC003058	AC003058 Arabidops
40	92	4.7	325	6	AX090357	AX090357 Sequence
41	88.4	4.6	396	6	AX045795	AX045795 Sequence
42	86.2	4.4	262797	3	CNS07EGG	AL590449 chromosom
43	78.4	4.0	185932	2	AP003714	AP003714 Oryza sat
44	75.2	3.9	3024	10	CGU47320	U47320 Cricetulus
45	71.6	3.7	40494	2	AC005122	AC005122 Drosophil

ALIGNMENTS

RESULT 1	AF298815	AF298815	1964 bp	mRNA	linear	PLN 16-OCT-2000
LOCUS	Perilla frutescens diacylglycerol acyltransferase (Dgat1) mRNA,					
DEFINITION	complete cds.					
ACCESSION	AF298815					
VERSION	AF298815.1	GI:10803052				
KEYWORDS						
SOURCE	Perilla frutescens.					
ORGANISM	Perilla frutescens					
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
	Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.					
	1 (bases) to 1964)					
	Hwang,S.-K. and Hwang,Y.-S.					

Isolation of *Perilla frutescens* diacylglycerol acyltransferase cDNA

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1964)  
AUTHORS Hwang, S.-K. and Hwang, Y.-S.  
TITLE Direct Submission

JOURNAL Submitted (23-AUG-2000) Division of Biochemistry, NIAST, RDA, 249  
Seodun-dong, Gwonsun-gu, Suwon 441-707, Korea

## FEATURES

## source

Location/Qualifiers

1..1964

/organism="Perilla frutescens"

/cultivar="Okdong"

/db\_xref="taxon:48386"

1..1964

/gene="Dgat1"

69..1673

/gene="Dgat1"

/note="DGAT"

/codon\_start=1

/product="diacylglycerol acyltransferase"

/protein\_id="AAG33696.1"

/db\_xref="GI:10803053"

/translation="MAILDSPEILDITSSADNGAAHHTLRRROSARSVPPLDSDS

NSLAEASINDSENVNRDANLLENLGGAVESENEKOESEYKKEGAKVKEGFTSGN

GTDMVAKFTFPPAPAPHRKNKESPLSSDAIFKQSHAGLFLNLCIVLVAVNSRLIEN

LMKGNLKSQFNSSTSLRDLMLCCLSLPVALASFLVEKLVKLNLYIPENWAVEL

HVTITVEILFPVVVILURCDASVLSGVTLMFACVTWKLVSXHTNYDLRLVLSLD

KWEAMSYWLDAYDYSPKSLAYFMVAPITLCQPSYPRACIRKGVVVRQLIKLVIF

TGLMGFTIEQYINPVGNSQHPKLNLYAIERVLKLSVFNLYVWLCMFYCFPHLWN

ILABLFCGDFEYKQNNARTVEEYWRMNPVHKWVRHLYPCPLQNGIPKIVAVL

IARLVSAIFHELCAVPCQIFKFWAFSGIMLQVPLVITVNYLQEFKNSVGNMFNC

FFCIEGPGMVCVLLYHGLDMNRKASAR"

BASE COUNT 522 a 395 c 461 g 586 t

## ORIGIN

Query Match 41.7%; Score 810.2; DB 8; Length 1964;

Best Local Similarity 76.8%; Pred. No. 1.3e-176;

Matches 989; Conservative 0; Mismatches 298; Indels 0; Gaps 0;

QY 243 AAGACACTGATTTCTCCCTCCTCAAAATTCGGCTACCGCTCCCTCCGTCGCCGCTCACCGCA 302  
DB 373 ACGGAAGTATGTTTGGCGCTCAAAATTCACATTCAGGCGCGCGCGCTGCTCACCGCA 432  
QY 303 AAGTGAAGAAAGCCGCTCAGCTCCGACACTATTTCCGTCAGAGTCACGCGGCGCTCT 362  
DB 433 AAAATAAGGAGAGTCTCTTACGCTCCGACGCTCTTCAAAACAGAGCATCAGCGCTCT 492  
QY 363 TCAACTTTGTATAGTACTGCTTGTGCTGTAAGTAGCGACTCATCATTCAGAAATTTAA 422  
DB 493 TCAACTTTGTATAGTACTGCTTGTGCTGTAAGTAGCGACTCATCATTCAGAAATTTAA 552  
QY 423 TCAAGTATGTTGTTGATGATCAAAATTCGGCTTTTGGTTTGTAGTTCAAGATCATTCAGAGACT 482  
DB 553 TGAAGTATGTTGTTGATGATCAAAATTCAGGATTTTGGTTTGTAGTTCAACATCGCTAGGATT 612  
QY 483 GGCCCTTTTCATGTTGTTCTTCTCTTGTGTTGTTATTTCCCTTTCGCTGCCCTTTATATGG 542  
DB 613 GGCCACTCTAATGTTGTTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 672  
QY 543 AGAAGTGGCAACGGAAGTATACCCGACAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 602  
DB 673 AGAAGTGGCAACGGAAGTATATATACCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 732  
QY 603 TTACCTCAACTTCGCTTTTCATCAGTTTATGTTTATTTCTCAGGTGTTGTTGTTGTTGTTGTT 662  
DB 733 TCACAACAGTGGAAATCTGTTTCCAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 792  
QY 663 TATCAGGTGTCAGTTTATGTTGTTTCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 722  
DB 793 TATCAGGTGTCAGTTTATGTTGTTTCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 852  
QY 723 CACATACAAATCATGATGAGGACACTTACCAAAATAGTTTGAAGGAGGAGGAGGAGGAGGAGG 782  
DB 853 CACATACAAATCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 912

QY 783 TCGATACCTCTGAACATGAGCTATCCTTACAAAGCTAAGCTTCAAGAGCTTGGCATATTTCC 842  
DB 913 CCAGGTACTGGAACCTCGACTACGCTTATGATGAAGCTTTAAGAGCTCGGCATATTCA 972  
QY 843 TGGTGGCCCTTACATATGTTTACCAGCAAGCTATCTCCGACACCTTATATTCGAAGG 902  
DB 973 TGGTGGCTCTACATGTTTACCAGCAAGCTATCTCCGACCAAGCTTACCTCGGACCTTGG 1032  
QY 903 GTTGGTCTTTCGCAACTTCTCAAGCTGATAATATTTACAGGAGTATGGATTTATAA 962  
DB 1033 GTTGGTGGTGAAGCAACTAATTAAGCTGGTAATATTCACAGGACTCATGGATTTATTA 1092  
QY 963 TAGACCAATATTAATCCCATAGTACAAAATTCACAGCATCCTCTCAAGGAAACCTTC 1022  
DB 1093 TAGAACAGTACATAAACCCGATCGTTCAAAATTCACACATCCTCTGAAAGGAAACCTTT 1152  
QY 1023 TTTAGCCACCGAGAGAGTCTCGAAGCTTTCTGTTTCCAAATTTATATGTGGCTCTGCA 1082  
DB 1153 TATATGCCATTTGAGAGGGTCTTGAAGCTTTCTGTTTCCAAATTTATATGTGGCTCTGCA 1212  
QY 1083 TGTCTATTGCTTTTCCACCTTTGGTTAAATATCTCCGACAGCTTCTTCGATTTGGTG 1142  
DB 1213 TGTCTATTGCTTTTCCACCTCTGGCTAAATATCTTCTGTAACCTCTCTGCTGCTTTGGGG 1272  
QY 1143 ATCGTGAATCTACAAGGATTTGGTGAATGCCAAACTGTGCAAGATTTATTTGGAGGATGT 1202  
DB 1273 ACCGTGAGTTTATAAGGATTTGGTGAATGCCAGACAGTGGAGAGTACTGGAGAATGT 1332  
QY 1203 GGAATATGCTGTTTCACAAATGATGATGCCGACCTATATATTTCCATGTTTAAAGGACG 1262  
DB 1333 GGAATATGCTGTTTCACAAATGATGATGTTGTTGCGCATATATATCTGCTCCATGTTCA 1392  
QY 1263 GTCTACCAAAAGCTGCTGCTCTTTTAAATTTGCTTCTGCTTCTGCTTATTTCCATGAGC 1322  
DB 1393 GAATACCAAAAGTATGTCGAGCTTTTGTGCTGCTGCGGATTTTTCATGAGC 1452  
QY 1323 TGTGATCTGTTTCTTCCCTTCCACATATTTCAAGTTTGGGGCTTTCCGGTGAATTTATTTTC 1382  
DB 1453 TGTGCTGTCAGTCCCTTGGCCAAATATTTCAAGTTTGGGGCTTCTCGGATCATGCTTC 1512  
QY 1383 AGGTCTCTTGGTCTGATGATCACTAATTTATCTGCAAAATATAATTCAGAACTCAATGGTTG 1442  
DB 1513 AGGTCTCTGATCTGATGATCACTAATTTATCTGCAAAATATAATTCAGAACTCAATGGTTG 1572  
QY 1443 GAAATATGATTTTGGTTCATATTTCACTATCTTGGTCAACCTATGTTGCTACTGCTAT 1502  
DB 1573 GCAATATGATTTCTGTTGCTTCTTCTGATCTTTGGTCAACCTATGTTGCTGTTGCTGT 1632  
QY 1503 ACTACCATGACTTGATGAATAGGAAAG 1529  
DB 1633 ACTACCATGACTTGATGAATAGGAAAG 1659

## RESULT 2

## AY084052

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## 2 (bases 1 to 2090)

## Tropaeolum majus

## Tropaeolum majus

## Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

## Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

## Rosidae; eurosids II; Brassicales; Tropaeolaceae; Tropaeolum.

## Mietkiewska, E., Pedersen, K., Katakic, V. and Taylor, D.C.

## Characterization of a putative diacylglycerol acyltransferase mRNA

## from Tropaeolum majus embryo

## Unpublished

## 2 (bases 1 to 2090)

AY084052 2090 bp mRNA linear PLN 08-APR-2002  
Tropaeolum majus putative diacylglycerol acyltransferase mRNA,  
complete cds.  
AY084052  
AY084052.1 GI:20086323

Tropaeolum majus.  
Tropaeolum majus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Tropaeolaceae; Tropaeolum.  
Mietkiewska, E., Pedersen, K., Katakic, V. and Taylor, D.C.  
Characterization of a putative diacylglycerol acyltransferase mRNA  
from Tropaeolum majus embryo  
Unpublished  
2 (bases 1 to 2090)

AUTHORS Mietkiewska, E., Pedersen, K., Katavic, V. and Taylor, D.C.  
TITLE Direct Submission  
JOURNAL Submitted (12-MAR-2002) Seed Oil Biotechnology, National Research Council of Canada, Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada  
FEATURES Location/Qualifiers  
source 1. 2090  
/organism="Tropaeolum majus"  
/db\_xref="taxon:4020"  
171..1727  
CDS  
/product="putative diacylglycerol acyltransferase"  
/protein\_id="AA003340.1"  
/db\_xref="GI:20086324"  
/translation="MAVAESQNTMTTMSGHSDLNFRFRKPPSSVIEPSSSGFTST  
NPVATGVRHNRDQDRVGAENATGCVNLGNQGVVIGNEEKQVGETDIFRTYRPS  
FPAHRVRESPLSDDAIKQSHAGLNLICIVVLIANVNSRLIENLMKYGLWLDGFWF  
NSRGLDWSIECCCLTPIPLAAFTVEKLVORNIHSELVALLHVIVSTAALYPIV  
VLTCDVYMSGVLELFCGIMWLKLYSAHTSSDIRLAKSGYKGDHPNSTIVSCS  
YDVLSKLAIFWVATLCQPSYRSCIRKGVVQFVKLIVIGLMGFITIEQINP  
IVRNSKHPKLGDFLXAIERVLKSPNLYVWLCMFYSFPHLWNLIAELLRFGDREFY  
KDWNAKTVAEYFWKMMNPFVHRWVHLVFPCLRNIGIPKEGAILIAFLVSGAFHELIC  
AVPCHVFKLWAFIGIMFQVPLVLIITNYLQEKFSNVMGNMIEFWIFCILGQPMCVLLY  
YHDLNLNKEK"  
BASE COUNT 527 a 397 c 458 g 708 t  
ORIGIN  
Query Match 38.8%; Score 754; DB 8; Length 2090;  
Best Local Similarity 74.7%; Pred. No. 1.2e-163;  
Matches 946; Conservative 0; Mismatches 320; Indels 0; Gaps 0;  
QY 268 ATTGGCTACCGCTTCCTCGTCCCGCTCACCAGGAAGTGAAGGCGCTCAGCTC 327  
DB 458 ATTCACCTTACCGGCTTCGTTCCGCTCATCGGAGGGTGAGGAGAGTCTCTAGCTC 517  
QY 328 CGACACTATTTCCTCGTCCAGAGTCAACGGCCCTCTCAACCTTGTATAGTACTGCTGT 367  
DB 518 TGATGAATCTTCAACAGAGCCATCGGGTTTATCAACCTGTGTATAGTACTGCTCAT 577  
QY 388 TCGTGTGAATPAGCGACTCATATTGAGAAATTAATGAAGTATGGTTGGTTGATCAAAATC 447  
DB 578 TGCAGTAAACAGTAGCTTATCATCGAAAATCTTATGAAGTATGGTTGGTTGATCGATAC 637  
QY 448 TGCCTTTGGTTAGTTCAAGTCANTGAGAGCTGGCCCTTTTCATGTGTTGCTCTTC 507  
DB 638 TGGTTTCTGCTTAACTCAAGATCACTGGGTGATGCTCATCTTTATGCTGCTGTAC 697  
QY 508 TCTTGTGTTATTCCTTCCTGCTCTTTATAGTGGAGAAGTGGCACACGGGAAGTGTAT 567  
DB 698 ACTCCCAATTTCCCACTGCTGCTTTATGTTGAAAGCTGGTGAGGAAATCATAT 757  
QY 568 ACCCGAACCAAGTTGTTGTTGTTACTTCATATAATCATCACTCAACTTCGCTTTTCTATCC 627  
DB 758 ATCTGAACCTGTTGCTTCTCTCATGTAATCGTTTCTACCGCTGCACTTTTATATCC 817  
QY 628 AGTTTATGTTTCTCAGGTGATCTGCTTTGTTGATCAGGTGTCACGTTAATGCTGTT 687  
DB 818 AGTTATGTTGATCTTAACGTGTATTCGGTGTATATGCTGCTGTGTTGTTGATGCTCTT 877  
QY 688 TCTCTGTTGTTATGTTAAATTTGTTGCTTATGACATACAAATCATGATATGAGAGC 747  
DB 878 TGGTTGATATGTTGTTGAAGCTGGTGTGATGACATACATAGTTCTGATATTAGAAC 937  
QY 748 ACTTACCAATATGTTGAAAGCGAGAACCACTGCTCGATACATCTGAACATGACTATCC 807  
DB 938 ACTGGCAATCTGGCTATAAGGGGATCGCACCCCAATCAACCATTTGAGTTGCTC 997  
QY 808 TTACACGTTAAGCTTCAAGAGCTTGGCATATTTCTCGTTGGCCCTACATATGTTACCA 867  
DB 998 ATATGATGTTAGCTTGAAGAGTTTGGCATACTTCACTGTTGGCGCCGACATTTGTTACCA 1057  
QY 868 GCCAAGCTATCCCGCACACCTTATATTCGAAAGGCTTGGTTGTTTCGCCCACTTGCTCAA 927  
|||||

DB 1058 GCTAGCTATCTCGTCTCGTGTATCCGCAAGGTTGGGTTGTCGAATTTGTCAA 1117  
QY 928 GCTGATAAATTTACAGAGTTATGGATTTATAATAGACCAATATATAATATCCCATAGT 987  
DB 1118 ACTAATAGTTTTTCATAGCTCATGGGTTTCAATAGAACAAATATATAATCTATCGT 1177  
QY 988 ACNAAATTTACAGCTCTCTCAAGGAAACCTTCTTTAGCCCAACGAGAGATTTCTGAA 1047  
DB 1178 TCGAAATTTCAACACCCCATTTGAAAGGAGATTTTATATGCAATAGAAAGATTTTGA 1237  
QY 1048 GCTTTCTGTTCCAAATTTATATGTTGCTCTGCTATGTTCTTATGTTTTTCCACCTTTG 1107  
DB 1238 GCTTTCAGTTCCAAATCTATATGTTGGCTTTCGATGTTCTACTCTTTTCCACCTCTG 1297  
QY 1108 GTTAAATATCTCTGGCAGAGCTTCTTCGATTTGGTATCGTGAATTTACAAGGATTTGGT 1167  
DB 1298 GTTGAACATACTGGCTGAGCTTCTTCGCTTTGGTGTATCGTGAATTTACAAGATTTGGT 1357  
QY 1168 GAATGCCAAACTGTCGAAGATTTATTGGAGGATGTGGAATATGCCCTGTTCAAAATGGAT 1227  
DB 1358 GAATGCCAAACTGTTGCGAGATTTTGGAAATGTGGAATATGCCCTGTTTCATAGATGGAT 1417  
QY 1228 GATCCGCCACCTATATTTTCCATGTTTAAAGGCACGCTTACCAAGGCTGCTCTCTTT 1287  
DB 1418 GGTTCGTCATCTATATTTTCCCTGTTTGGAGAAATGGATACCAAGGATTTGCCATTAT 1477  
QY 1288 AATTCGCTTCTGCTTCTGCTTTTATTCATGAGCTGTGCTGCTTCTTCCCTTGGCCACAT 1347  
DB 1478 TATCGGCTTCTAGTTTCTGCTGCTTTCCATGAGCTTGCATTCGATTTCCCTTGGCCACGT 1537  
QY 1348 ATTCAGTTGTTGGCTTTCGCTGGAATTTATGTTTCAGTTTCCCTTGGCTTGTATCACTAA 1407  
DB 1538 ATTCAGTTTATGGCTTCTTATAGGCATTTATGTTTCAGTTTCCCTTGGTATTTACGAA 1597  
QY 1408 TTATCTGCAAAATAAATTCAGAACTCAATGTTGGAATATGATTTTGGTTCATATT 1467  
DB 1598 TTATCTACAGAAAGTTTCAGTATTTATGTTGGGCAATATGATCTTCTGGTTATCTT 1657  
QY 1468 CAGTATCTTGGTCAACCTATGTTGCTACTGCTATCTACTACCATGACTTGTGATGAATAGAA 1527  
DB 1658 CTGCACTATGGCCAACTTATGTTGCTCTTCTATATATACCATGACCTGATAAATCTAAA 1717  
QY 1528 AGGCAA 1533  
DB 1718 GGAANA 1723  
RESULT 3  
AF129003  
LOCUS  
DEFINITION Nicotiana tabacum diacylglycerol acylCoA acyltransferase (DAGAT)  
mRNA, complete cds.  
ACCESSION AF129003  
VERSION AF129003.1 GI:6625652  
KEYWORDS  
SOURCE Nicotiana tabacum.  
ORGANISM Nicotiana tabacum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
REFERENCE 1 (bases 1 to 2099)  
AUTHORS Bouvier-Nave, P., Benveniste, P., Oelkers, P., Sturley, S.L. and Schaller, H.  
Expression in yeast and tobacco of plant cDNAs encoding acyl CoA:diacylglycerol acyltransferase  
Eur. J. Biochem. 267 (1), 85-95 (2000)  
JOURNAL MEDLINE 20069349  
PUBMED 10601854  
REFERENCE 2 (bases 1 to 2099)  
AUTHORS Benveniste, P.  
Direct Submission  
TITLE Submitted (17-FEB-1999) Plant Molecular Biology Institute, C.N.R.S., 28 rue Goethe, Strasbourg 67083, France

Query Match	38.3%	Score 744..6;	DB 8;	Length 2099;
Best Local Similarity	73.9%	Pred. No. 1.9e-161;		
Matches 964;	Conservative 0;	Mismatches 329;	Indels 12;	Gaps 1;
Qy	228	AGCAAAACGAAAAACAAGACACTGATTTCTCCGTCCTCAAAATCGCCCTACCGCTCCCTCCG	287	
Db	503	AGNAACCAATCAACGGAACAACAAGACGCCGTTTAAATTTGCTTACAGGCGCTCG	562	
Qy	288	TCGCCGCTCACCGCAAAAGTGAAGGAAAGTCGCCGTACGCTCCGACACTATTTTCCGTCGAGA	347	
Db	563	CACCAGCTCACCGCGCAATCAAGGAGAGTCTCTCAGCTCCGAGGCCATTTTCAACAGA	622	
Qy	348	GTACCGCGGCCCTCTCAACCTTTGTATAGTAGTCCTTGTCTGTGTAATAGCCCACTCA	407	
Db	623	GTACCGCAGCGCTGTTCAATCTCTGTGTGGTGGCTGATGTGTTTAAACAGCAGGCTGA	682	
Qy	408	TCATTGAGAAATTAATGAAGTATGGTTGGTTGATCAAAATCTGGCTTTTGGTTTGAATCAA	467	
Db	683	TTATCGAGAACTTGATCAAGTATGGCCCTTTTAATTAGGCTGGCTTTTGGTTTGAATCGA	742	
Qy	468	AGTCATTGAGAGACTGCCCCCTTTTCAATGTTGTGTCTTCTCTGTGTGGTAATTTCCCTTCG	527	
Db	743	AGTCGTTGAGGGATGTGCCCGCTTCTAATGTGTCTCTCAGTCTCTCAAAATTTTGGCGCTCG	802	
Qy	528	CTGCCCTTTATAGTGGAGAAGTTGGCACACGGAAGTCTATACCGCAACCACTGTTCTTCTTG	587	
Db	803	CTGCTTTTCTTGTGGAGAAGTTGGCACACGAGAGGCAATTTGACGAGCGTGGCGTGTTA	862	
Qy	588	TACTTCATAATCATTTACCTCAACTTCGCTTTTCTATCCAGTTTATAGTTATTTCTCAGGT	647	
Db	863	CTCTTCACATAACTATTAACAACAGCTGCCCATTTGTATCCAGTTCTGGTCATTTCTTGGGT	922	
Qy	648	GTGATTCGCTTTGTATACAGGTGTACAGTAAATGCTTTTCTTGTGTGTGTATAGGTTAA	707	
Db	923	GTGATTCGCTTTTCTGTTGGTGTCAATATGATGCTGGTGTGCTTGCATTTGTGTGGATGA	982	
Qy	708	AATTGTGCTCTTATGACACATACAAACTATGATATGAGAGCACTTACCAAAATTAGTCAAA	767	
Db	983	AGCTGGTTTCTTAGGCACATACAAATCATGATATGACAGAGCTCGCAAGCTCTACGGACA	1042	
Qy	768	AGGAGAAACACTGCTCGATCTCTGAAACATGGACTATCCTTACAACGTAAGCTTCAAGA	827	
Db	1043	AGGATGAA-----ACCTCAGATGGGATTTCTCTTATGATGTGTAGCTTCAAGA	1090	

source 1. .1904  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
gene 139. .1701  
/gene="Dgat"  
CDS 139. .1701  
/gene="Dgat"  
/function="lipid synthesis"  
/codon\_start=1  
/product="diacylglycerol acyltransferase"  
/protein\_id="CAB45373.1"  
/db\_xref="GI:5123718"  
/translation="MAILDSAGVTVTENGCGEFVDLRLRRKSRSDSSNGLLLSGS  
DNNSPSDDVADPADVRDIDSVDNDADGTANLADNNGGDDNNGGGRGGGRGNAD  
AFTTYRSPVAGRAESESPLVDSDAIFKQSHAGLFNLVYVLIANVRLIENLMRYGW  
LIRDFWSSRLSDPWFMCISLIFPLAAFTVEKLVQYISSEPVVIFHIIITM  
TEVLVYVTLRCDASFLSGVTMLLTCTIVWLKLSVYAHTSYDISELANADKANPEV  
SYVLSKSLAYPMVAPTLCYOPSPRSACIRKGMVAROPAKLVIETGPMGFLLILOYIN  
PIVRNKHPLGKDLIXATERVILKLSVPLNLYVWLMPYCFPHLWNLILAEILLCFGDBREF  
YKDMNNAKSGDIWRMNNPVHKMVRHYIFPLKRSKIPKTLAILIILAEILSVAFVHELIC  
IAVPCRLFKLMAFLGIMFQVPLVFTTNYLQERFSGTVGNMIFWFIICIFGQPMCVLLY  
YHDLNMRKGSMS"

BASE COUNT 458 a 399 c 431 g 616 t  
ORIGIN

Query Match 37.1%; Score 719.8; DB 8; Length 1904;  
Best Local Similarity 73.2%; Pred. No. le-155;  
Matches 960; Conservative 0; Mismatches 327; Indels 24; Gaps 2;

QY 241 ACAAGACACGATTTCTCGCTCCTCAAAATTCGGCTACCGTCCTCGCTCCCGCTCACCG 300  
DB 423 AGAAGNAGGAAAGCCGATGCTAGCTTACGTATCGACCGCTCGGTCACGCTCATCG 482  
QY 301 CAAAGTGAAGGAAGTCCGCTCAGCTCGGACACTATTTCCGTCAGAGTCACGCGGCGCT 360  
DB 483 GAGGCGGAGAGAGTCCACTTACGCTCGGACCAATCTTCAACACAGAGCCATCGCGATT 542  
QY 361 CTTCAACCTTTGTATAGTAGTCTTGTGCTGTAATAGCCGACTCATCTTCAGATTT 420  
DB 543 ATTCAACCTCTGTGTAGTCTTATAGCTGTAAACAGTAGACTCATCATCGAANAATCT 602  
QY 421 AATCAAGTAGTGGTTGTATCAAACTCTGGCTTTTGGTTTGTAGTCAAGTCATTGAGAG 480  
DB 603 TATGAGTAGTGGTTGTATCAGACGGAATTTCTGGTTTGTAGTCAAGATCGCTCGGAGA 662  
QY 481 CTGCGCCCTTTTCATGTGTCTTTCTCTCTGTGGTATTTCCCTTCGCTGCCCTTTATAGT 540  
DB 663 TTGCGCGCTTTTCATGTGTGTATATCCCTTTTCGATCTTTTCCCTTGGCTGCCCTTACGGT 722  
QY 541 GGAGAGTTGGCACAAGGAAGTGTATACCCGAACAGTTGTTGTGTACTTCAATATAT 600  
DB 723 TGAGAAATTGGTACTTCAGAAATACATATCAGAACCTGTGTGATCTTTCTTTCATATAT 782  
QY 601 CATTACTCAACTTCGCTTTCTATCCAGTTTGTAGTTATTTCTCAGGTGTGATTTCTGCTTT 660  
DB 783 TATCAGATGACAGAGTTTGTATCCAGTTTACGTCACCCCTAAGGTGTGATTTCTGCTTT 842  
QY 661 TGTATCAGGTGTACAGTTTATGCTGTTTCTTGTGTGTATGTTTAAATTTGGTGTCTTA 720  
DB 843 TTTTATCAGGTGTACATTTTGTATGCTCTCACTTGCATTTGTGTGCTTAAAGTTGTTCTTA 902  
QY 721 TGACATACAAACTATGATGAGAGCACTTACCAATTTAGTTTGAAGGGAGAGCACT 780  
DB 903 TGCTCATACTAGTATGACATAAGATCCCTTAGCCAAATGCAAGCTGATAAGG----- 952  
QY 781 GCTCGATACTCTCAACATGGACATCTTCAACAGTAAAGCTTCAAGAGCTTGCATATTTT 840  
DB 953 -----CCAACTCGAATCTCTACTAGTACGTTTCAAGAGCTTGGCATATTT 1001  
QY 841 CTTGGTTGCCCTTACATTTATGTTTACGAGCAAGCTATCTCTCGCACACCTTTATATCGAAA 900  
DB 1002 CATGCTCGCTCCACATCTGTTTATCAGCAAGTTATCCAGCTTCTCCATGTATACGGAA 1061

QY 901 GGTTGGTTGTTTCCGCCAACTTGTCAAGCTGATAATATTTACAGGAGTGTATGGATTTAT 960  
DB 1062 GGTTGGTTGGCTGCTGCAATTTGCAAAACTGGTCAATATTCACCGGATTCATGGGATTTAT 1121  
QY 961 AATAGACCAATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGGAACCT 1020  
DB 1122 AATAGAACCAATATATAATCCCTATTGTTCAGGAACCTCAAAGCATCTCTTGAAGGCGATCT 1181  
QY 1021 TCTTTACGCCACCGAGAGAGTCTTGAAGCTTTTCTGTTCACAAATTTATATGTGTGGCTCTG 1080  
DB 1182 TCTATATGCTATTGAAAGAGTGTGAAGCTTTTCAGTTCCAAAATTTATATGTGTGGCTCTG 1241  
QY 1081 CATGTTCTATGCTTTTCCACCTTTGTAATATATCTTGGGAGAGCTTCTTCGATTTGG 1140  
DB 1242 CATGTTCTACTGCTTCTCCACCTTTGGTAAACATATATGGCAGAGCTCTCTGCTCGG 1301  
QY 1141 TGATCTGGAATCTCAAGAGATTGGTGAATGCCAAACTGTGCAAGATTATTTGGAGGAT 1200  
DB 1302 GGATCTGGAATCTCAAGAGATTGGTGAATGCCAAAGTGTGGAGATTACTTGGAGAAAT 1361  
QY 1201 GTGGAATATGCCTGTTTCACAAATGATGATCCGCCACCTATATTTCCATGTTTAAAGGCA 1260  
DB 1362 GTGGAATATGCCTGTTTCATAAATGATGATGTTTCGACATATATATCTCCCGTGTGCGCAG 1421  
QY 1261 CGTCTACCAAGGCTGCTCTCTTTTAATGCTTCCCTGGTTTCTGCTTTATTCATGA 1320  
DB 1422 CAAAGTACCAAGACACACCTGCGCATTTATCATTTGCTTCTCCTAGTCTCTGCACTTTTCA 1481  
QY 1321 GCTGTGCTGTTGTTTCCCTGTCACATATTTCAAGTTGTGGCTTTCGGTGGAAATTTATGTT 1380  
DB 1482 GCTATGCTGCGAGTTCCTTGTGCTCTCTTCAAGCTATGGCTTTTCTTGGGATTTATGTT 1541  
QY 1381 TCAGGTTCCTTTGGTCTTGATCACTAATATCTGCAAAATATAATTCAGAACTCAATGTT 1440  
DB 1542 TCAGGTGCCTTTGGTCTTCATCAACAACTATCTACAGGAAGGTTTGG---CTCAACGGT 1598  
QY 1441 TGAATATGATTTTGGTTCATATTCAGTATCTTGGTCAACCTATGTGTGTACTGCT 1500  
DB 1599 GGGGAACATGATCTTCTGGTTTCTCTCTGATTTTCGGACAAACCGATGTGTGCTTCT 1658  
QY 1501 ATACTACCATGACTTGATGAATAGGAAGCAAACTTGACTGAAGCTACGG 1551  
DB 1659 TTATTACCAGCACTGATGAAACCGAAAGGATCGATGTCATGAACAACATG 1709

RESULT 5  
LOCUS AX090345  
DEFINITION Sequence 38 from Patent WO0116308.  
ACCESSION AX090345  
VERSION AX090345.1 GI:13444207  
KEYWORDS  
SOURCE thale cress.

ORGANISM  
Arabidopsis thaliana  
Bakryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE  
AUTHORS Lasser M. and van Eenennaam A.  
TITLE Plant sterol acyltransferases  
JOURNAL Patent: WO 0116308-A 38 08-MAR-2001;  
MONSANTO COMPANY (US)

FEATURES  
source  
1. .1942  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"

BASE COUNT 461 a 421 c 425 g 635 t  
ORIGIN

Query Match 37.1%; Score 719.8; DB 6; Length 1942;  
Best Local Similarity 73.2%; Pred. No. le-155;  
Matches 960; Conservative 0; Mismatches 327; Indels 24; Gaps 2;



QY 241 ACAAGACACTGATTTCTCGTCCCTCAAAATTCGGCTACCGTTCCTCCGTCCTCCGCTCACCG 300  
Db 521 AGAAGAGAGAGAAACGCGGCTACGTTTACGTTATCGAGTCGGTTCACGACTCATCG 580  
QY 301 CAAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTTCGGTCAGAGTACGCGGGCT 360  
Db 581 GAGGCGGAGAGAGTCCACITAGCTCCGACGCAATCTTCAACAGAGCATCGCGGAT 640  
QY 361 CTTCAACCTTTGTATAGTAGTCTTTTGTGCTGTGAATACCGCACTCATCATTTGAGAATTT 420  
Db 641 ATTCAACCTCTGTAGTAGTCTTTTGTGCTGTGAATACCGCACTCATCATTTGAGAATTT 700  
QY 421 ATGAAGTATGTTGGTGTGATCAAAATCTGGCTTTAGTTTAAAGTCAATTTAGT 540  
Db 701 TATGAAGTATGTTGGTGTGATCAAAATCTGGCTTTAGTTTAAAGTCAATTTAGT 540  
QY 481 CTGGCCCTTTTCATGTGTGCTTTCTCTGTGCTATTTCTCTTCGCTCCGCTTATAGT 540  
Db 761 TTGGCCGCTTTTCATGTGTGCTTTCTCTGTGCTATTTCTCTTCGCTCCGCTTATAGT 540  
QY 541 GGAGAAAGTTGGCACAACGGAAGTGTATACCCGAACCACTTGTGTGCTATCATATAAT 600  
Db 821 TCAGAAATTTGGTACTTCAGAAATACATATCAGAACCTGTGTGCTATCATATAAT 880  
QY 601 CATTAACCTCAACTTCGCTTTTCTATCCAGTTTGTAGTTATCTCAGGTGTGATCTGCTTT 660  
Db 881 TATCACCATGACAGAGGTTTGTATCCAGTTTATACCTACCCCTAAGGTGTGATCTGCTTT 940  
QY 661 TGTATCAGGTGTCAGTTAAAGTCTGTTTCTGTGCTGTGCTTAAAGTGTGCTCTCTTA 720  
Db 941 TTTATCAGGTGTCAGTTAAAGTCTGTTTCTGTGCTGTGCTTAAAGTGTGCTCTCTTA 1000  
QY 721 TGCACATACAAACTATGATATGAGACACTTTACCAAAATAGTTTGAAGGAGGAGCACT 780  
Db 1001 TGCTCATACTAGTATGACATGAAGATCCCTAGCCCAATCGACGTGATAAG 1050  
QY 781 GCTCATACTCTGAACATGGACTATCTTACACCTAGTCTTCAAGACTTGGCATATTT 840  
Db 1051 -----CCAACTCTGAAGTCTCTCTACTACGTTTGAAGAGCTTGGCATATTT 1099  
QY 841 CCGTGTTCGCCCTACATATGTTACACGCAAGCTATCTCGCACACTTATATTCGAAA 900  
Db 1100 CATGTCGCTCCACATTTGTTATCAGCCCAAGTTATCCAGCTTCTGATGTATACGAA 1159  
QY 901 GGGTGGTGTGTTCCCAACTTGTCAAGCTGATATATTTACAGAGTATGGGATTTAT 960  
Db 1160 GGGTGGTGGTCTGCTCAATTTGCAAACTGTGTATATTCACCGCATTCATGGGATTTAT 1219  
QY 961 AATAGACCAATATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGGAACCT 1020  
Db 1220 AATAGAACAAATATATAATTCCTATTTGTCAGGAACCTCAAAAGCATCTCTCTGCTTCG 1279  
QY 1021 TCTTTACGCCACGAGAGTCTCAAGCTTCTGTTCCAAATTTATATGTTGGCTCTG 1080  
Db 1280 TCTATATGCTATTGAAGAGGTTTGAAGCTTTCAGTTCCAAATTTATATGTTGGCTCTG 1339  
QY 1081 CATGTTCTATTGCTTTTCCACCTTTGTTTAAATATCTCGCAGAGCTTCTTCGATTTGG 1140  
Db 1340 CATGTTCTATTGCTTTTCCACCTTTGTTTAAACATATTTGCGAGAGCTTCTCTGCTTCG 1399  
QY 1141 TGATCGTGAATTTCTACAGGATTTGTTGAATGCCCAGGCTTCTGCAAGATTTATGGAGAT 1200  
Db 1400 GATCGTGAATTTCTACAAAGATTTGTTGAATGCAAAAGATTTGCGAGATTTACTGGAGAT 1459  
QY 1201 GTGGAATATGCTGTTCACAAATGATGATCATCGCCACCTATATTTTCCATGTTTAAAGCA 1260  
Db 1460 GTGGAATATGCTGTTCATTAATGATGATGTTGCGATATATATCTTCCGCTGCTTCGCGAG 1519  
QY 1261 CGGTCTACCAAGGCTGCTGCTTTTAAATTCCTTCTGTTTCTGCTTTATTTCCATGA 1320  
Db 1520 CAAATACCAAAAGACACTCGGCATTTATCATTTGCTTTCTAGTCTCTGCGAGCTTTTCATGA 1579  
QY 1321 GCTGTGCATTTGCTGTTCTTCCCATGATATTCAGTTTGTGGGCTTTCGGTGGAAATATGTT 1380

Db 1580 GCTATGCATCGAGTTCCTTGTGCTCTTCAAGCTATGGGCTTTCTTGGGATATGTT 1639  
QY 1381 TCAGGTTCTCTTTGGTCTTGTGATCACTAAATATCTGCAAAATAAATTCAGAACTCAATGTT 1440  
Db 1640 TCAGGTGCTTTGGTCTTGTGATCACTCAAACTATCTACAGGAAGGTTTGG---CTCAACGGT 1696  
QY 1441 TGGAAATATGATTTTGTGTTTATATTCAGTATCTTGTCAACCTATGTGTGTAAGTCT 1500  
Db 1697 GGGGAACATGATCTTCTGTTTCTGATCTTTCGACAAACGATGTGTGCTTCT 1756  
QY 1501 ATACTACCATGACTGTGATGAATAGGAAGGCAAACTTACTGCAAGCTACGG 1551  
Db 1757 TTATTACCACGACTGATGAACCGAAAGGATCGATGTCATGAACAACTG 1807  
RESULT 6  
AX090349 AX090349 1942 bp DNA linear PAT 21-MAR-2001  
LOCUS Sequence 42 from Patent WO0116308.  
DEFINITION AX090349  
ACCESSION AX090349  
VERSION AX090349.1 GI:13444210  
KEYWORDS  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE  
1 (bases 1 to 1942)  
AUTHORS Luesner, M. and van Eenennaam, A.  
TITLE Plant sterol acyltransferases  
JOURNAL Patent: WO 0116308-A 42 08-MAR-2001;  
MONSANTO COMPANY (US)  
FEATURES  
Location/Qualifiers  
Source 1..1942  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
BASE COUNT 461 a 421 c 425 g 635 t  
ORIGIN  
Query Match 37.1%; Score 719.8; DB 6; Length 1942;  
Best Local Similarity 73.2%; Pred. No. le-155;  
Matches 960; Conservative 0; Mismatches 327; Indels 24; Gaps 2;  
QY 241 ACAAGACACTGATTTCTCGTCCCTCAAAATTCGGCTACCGTTCCTCCGTCCTCCGCTCACCG 300  
Db 521 AGAAGAGAGAGAAACGCGGATGCTACGTTTACGTTATCGAGTCGGTTCACGACTCATCG 580  
QY 301 CAAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTTCGGTCAGAGTACGCGGGCT 360  
Db 581 GAGGCGGAGAGAGTCCCACTTAGCTCCGACGCAATCTTCAACAGAGCCATGCGCGGAT 640  
QY 361 CTTCAACCTTTGTATAGTAGTCTTTTGTGCTGTGAATAGCGCACTCATCATTTGAGAATTT 420  
Db 641 ATTCAACCTCTGTAGTAGTCTTTTGTGCTGTGAATAGCGCACTCATCATTTGAGAATTT 700  
QY 421 AATGAAGTATGTTGGTGTGATCAAAATCTGGCTTTAGTTTAAAGTCAATTTAGT 480  
Db 701 TATGAAGTATGTTGGTGTGATCAAAACGATTTCTGGTTAGTTCAAGATCGCTCGGAGA 760  
QY 481 CTGGCCCTTTTCATGTGTGCTTTCTGTTGCTGATTTTCTTTCGCTGCTTTATAGT 540  
Db 761 TTGGCCGCTTTTCATGTGTGCTTTCTGTTGCTGATTTTCTTTCGCTGCTTTACGCT 820  
QY 541 GGAGAAAGTTGGCACAACGGAAGTGTATACCCGAACCACTTGTGTGCTATCATATAAT 600  
Db 821 TCAGAAATTTGGTACTTCAGAAATACATATCAGAACCTGTGTGCTATCATATAAT 880  
QY 601 CATTAACCTCAACTTCGCTTTTCTATCCAGTTTGTAGTTATCTCAGGTGTGATCTGCTTT 660  
Db 881 TATCACCATGACAGAGGTTTGTATCCAGTTTATACCTACCCCTAAGGTGTGATCTGCTTT 940  
QY 661 TGTATCAGGTGTCAGTTAAAGTCTGTTTCTGTGCTGTGCTTAAAGTGTGCTCTCTTA 720





Db 995 TCTCATACTAGTATGACATGAAGATCCCTAGCCAAATCAGCTGATAAGG----- 1044  
Qy 781 GCTGATGACTCTGAACATGAGACTATCTTACAACGTAAGCTTCAAGAGCTTGGCATATTT 840  
Db 1045 -----CCATCTCTGAAGTCTCTACTACTAGCTTGAAGAGCTTGGCATATTT 1093  
Qy 841 CCGTGTGCTCCCTACATATGTTACCAGCAAGCTATCCTCGCACACCTTATATTCGAAA 900  
Db 1094 CATGCTGCTCCACATATGTTATATCAGCAAGTATCCAGCTTCTGCATGTATACGGA 1153  
Qy 901 GGGTGTGTTGTTCCCAACTGTCAAGCTGATATATTTACAGAGATTTACGATTTAT 960  
Db 1154 GGGTGTGCTGCTCAATTTGCAAACTGCTCATATTTACCGGATTTATCGGGATTTAT 1213  
Qy 961 AATGACCAATATATATATCCATAGTACAAATTCACAGCATCTCTCAAGGGAACCT 1020  
Db 1214 AATGACCAATATATATATATCTTATGTCAGCAACTCAAGCATCTCTTGAAGGCGATCT 1273  
Qy 1021 TCTTTACCCACCGAGAGAGTCTGAAGCTTCTGTTCCAAATTTATATGTGTGCTCTG 1080  
Db 1274 TCTATATGCTATGAAGAGTGTGAAGCTTTCAGTTCCTCAAAATTTATATGTGTGCTCTG 1333  
Qy 1081 CATGTTCTATGCTTTTCCACCTTGTGTTAAATATCTGCGAGAGCTTCTTCGATTTGG 1140  
Db 1334 CATGTTCTATGCTTTTCCACCTTGTGTTAAACATATTTGCGAGAGCTTCTGCTTCCG 1393  
Qy 1141 TGATGCTGAATTTCTACAAGGATTTGTTGAATGCCAAATCTGCGAAGATTTATTTGGAGAT 1200  
Db 1394 GGATGCTGAATTTCTACAAGATTTGTTGAATGCCAAATCTGCGAAGATTTATTTGGAGAT 1453  
Qy 1201 GTGGAATATGCTGTTCACAAATGATGATCGCGACCTATATTTTCCATGTTTAAAGCA 1260  
Db 1454 GTGGAATATGCTGTTCACAAATGATGATCGCGACCTATATTTTCCATGTTTAAAGCA 1513  
Qy 1261 CGGTCTACCAAGGCTGCTGCTTTTAAATGCTGCTTCTGCTGCTTCTGCTTTATCCATGA 1320  
Db 1514 CAAGTACCAAGAGACCTGCGCATATATGCTTTTCTAGTCTCTGCGAGTCTTTCATGA 1573  
Qy 1321 GCTGTGATGCTGCTTCTGCTGATATTTCAAGTTGTGGCTTCTGCGGAAATATATTT 1380  
Db 1574 GCTATGATGCTGCTTCTGCTGATATTTCAAGTTGTGGCTTCTGCGGAAATATATTT 1633  
Qy 1381 TCAGTTGCTTGTGCTTGTGATCAATATCTGCAAAATTAATTCAGAACTCAATGGT 1440  
Db 1634 TCAGTTGCTTGTGCTTGTGATCAATATCTGCAAAATTAATTCAGAACTCAATGGT 1690  
Qy 1441 TGGAAATATGATTTTGTGTTTCAATTTCAATCTGCTGCTCAACCTATGCTGCTACTGCT 1500  
Db 1691 GGGNACATGATCTCTGCTGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTTCT 1750  
Qy 1501 ATACTACCATGCTGATGAATAGGAAAGGCAAACTTGGACTGAAGCTACGG 1551  
Db 1751 TTATTACCACGCTGATGAACCGAAAGGATGATGCTGATGCTGAACAACTG 1801

RESULT 8  
AY054480 2005 bp mRNA linear PLN 05-SEP-2001  
LOCUS Arabidopsis thaliana diacylglycerol O-acyltransferase (At2g19450;  
F3p11.5) mRNA, complete cds.  
DEFINITION  
ACCESSION AY054480  
VERSION AY054480.1 GI:15450799  
KEYWORDS FLI\_CDNA.  
SOURCE Arabidopsis thaliana.  
ORGANISM Arabidopsis thaliana  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 2005)  
Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M.,  
Palm,C.J., Bowser,L., Jones,T., Banh,J., Shinn,P., Carninci,P., Chen,H.,  
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,

Kawai,J., Kim,C., Lin,J., Liu,X., Li,X., Narusaka,M., Pham,P.K.,  
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,  
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.  
Direct Submission  
Submitted (28-AUG-2001) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN  
Arabidopsis Full-length cDNA"): Seki,M., Narusaka,M., Ishida,J.,  
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs: Nguyen,M.,  
Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,  
Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,  
Kim,C., Lin,J., Liu,X., Pham,P.K., Sakano,H., Shinn,P.,  
Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Southwick,A., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed  
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.  
(SSP/Stanford) contributed equally to this work as PIs.

Location/Qualifiers  
1. 2005  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/chromosome="II"  
/clone="RAFL07-14-014"  
/note="This clone is in pBluescript"  
1. 2005  
/gene="At2g19450; F3p11.5"  
162. 1725  
/gene="At2g19450; F3p11.5"  
/codon\_start=1  
/product="diacylglycerol O-acyltransferase"  
/protein\_id="AAK96671.1"  
/db\_xref="GI:15450800"  
translation="MAILDSAGVTVTENGGEFVDLRLRRKSRSDSSNGLLSGS  
DNPSDDVGPADVRDRIDSVYNDQAGTANLAGDNGDNGGGRGGGRGNAD  
ATITRPSVPAHRRARSPUSDAIFKQSHAGLNCVVLIAVLIANLIENLMKYGM  
LIRTFWFSRSLRDNPLFCCISLIFPLAFTVEKLVQKIVSEPVFLIITM  
TEVLYPVYTLRCDSPALSGVTLMLLTICIVLKLVSATSYDRIANAAADKANPEV  
SYVSKSLAYENVAFTPCQPSYACIRKGVARQAKLVFTGFMGFIETQIN  
PIVNSKHPKGLDLLVAIERVLKLSVPLNLYWLCMFYCFEHLNLIALLCIGDFREF  
YKDWNRKSVGDYWRMNPVHKWVRHIYFPCLRKIPKTLAIITAFILVSAVHELC  
IAYPCRLFLKWAFLGINFQVPLVITNYIQERFGSTVGNMIFWIFICIFGQPMCVLLY  
YHDLNMRKGSMS"

BASE COUNT 499 a 413 c 447 g 646 t  
ORIGIN

Query Match 37.1%; Score 719.8; DB 8; Length 2005;  
Best Local Similarity 73.2%; Pred. No. le-155;  
Matches 960; Conservative 0; Mismatches 327; Indels 24; Gaps 2;

Qy 241 ACAAGACACTGATTTCTCCGCTCCTCAAAATTCGCTACGCTTCTCCGCTCCGCTCACCG 300  
Db 446 AGAAGGAGAGAAACCGCGATGCTACGTTACGATCGACCGCTCGGTTCCAGCTCATCG 505  
Qy 301 CAAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTCCGTCAGAGTCAACGGGCGCT 360  
Db 506 GAGGCGAGAGAGAGTCCACTTAGCTCCGAGCAATCTTCAACAGAGCCATCGCGGATT 565  
Qy 361 CTTCAACCTTGTATAGTACGCTTGTGCTGAATAGCGGACTCATCATTTGAGAATTT 420  
Db 566 ATTCAACCTCTGTGTAGTAGTCTTATTTGCTGTAACAGTAGTACATCATCGAAAAATCT 625  
Qy 421 AATGAAGTATGTTGGTTGATCAAAATCTGGCTTTTGGTTTATGTTCAAGTCAATGAGAGA 480  
Db 626 TATCAAGTATGTTGGTTGATCACAACGGATTTCTGTTTCTGTTTACATGATCCCTGCCGAGA 685



QY	361	CTTCAACCTTTGTATAGTAGTCCTTGTGCTGTGAATACCCACATCATCATTTGAGAAATTT	420
DB	492	ATTCAACCTCTGTAGTAGTCTTATATGCTGTAAACAGACTAGACTCATCATCGAANAATCT	551
QY	421	AATCAAGTATGGTTGGTTGATCAAACTCTGGCTTTTGGTTTACTTCCAAAGTCAATTCAGAGA	480
DB	552	TATGAAGTATGGTTGGTTGATCAGACGGAATTCGTGGTTAGTTTCAAGATCGTCGCGAGA	611
QY	481	CTGGCCCTTTTCATGTGTTGTTCTTCTCTGTGTGTAATTCCTTCGCTGCGCTTTATAGT	540
DB	612	TTGGCCGCTTTTCATGTGTTGTTATATCCCTTTTCGATCTTTCCCTTTGGCTTTACGGT	671
QY	541	GGAGAAGTTGGCACAACGGGAAGTGATACCCGAACAGTTGTTGTATCTTCATATAAT	600
DB	672	TGAGAAATGGTACTTCAGAAATACATATACGAACCTGTGTGCATCTTCTTCATATAT	731
QY	601	CATTACCTCAACTTCGGCTTTTCTATCCAGTTTTAGTTATCTTCAGGTGTGATCTTCGCTTT	660
DB	732	TATCACCATGACAGAGTTTTGTATCCAGTTTACGTCACCCCTAAGGTGTGATCTTCGCTTT	791
QY	661	TGTATCAGGTGCACGTTAATGCTGTTTCTTCTGTGTGTATGGTTTAAATTTGGTGCTCTTA	720
DB	792	TTTATCAGGTCTCACTTTGATGCTCCTCACTTGGCATTTGTGGCTTAAAGTTGGTTCTCTTA	851
QY	721	TGCACATACAACATATCATATGAGACGACTTACCAAAATTAGTTGAAAAGGGAAGACACT	780
DB	852	TGCTCATACTAGTATGACATATAGATCCCTGACCAATGACGACTGATNAGG-----	901
QY	781	GCTCGATACTCTGAACATGGACATATCCTTTACAACGTAAGCTTCAAGAGCTTGGCATATTT	840
DB	902	-----CCAACTCTGAAGTCTCCTTACTAGTTAGCTTGAAGCTTGGCATATTT	950
QY	841	CCTGGTTGCCCTTACATATATGTTTACCAGCAAGCTATCCTGCGACACCTTATATTCGAAA	900
DB	951	CATGGTCGCTCCACATATGTGTTTATCAGCAAGATTATCCACGTTCTGCGATGATACGGAA	1010
QY	901	GGGTTGTTGTTTCGCAACTCTGCAAGCTGATAATATTTACAGCAGTTATGGGATTTAT	960
DB	1011	GGGTTGGTGGCTCGTCNATTTGCCAACTGGTCATATTCACCGGATTCATGGGATTTAT	1070
QY	961	AATAGACAATATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGGAAACCT	1020
DB	1071	AATAGAACAATATATAAATCCTATTCTCAGGAATCAAGACATCTTTGAAAGGCGATCT	1130
QY	1021	TCATTAGCCACGAGAGAGTCTGAAGCTTCTGTTCCAAATTTATATGTGGGCTG	1080
DB	1131	TCTATATGCTATTTCAAGAGCTGTGAAGCTTTTCAGTTCCAAATTTATATGTGGGCTG	1190
QY	1081	CATGTTCTATTGCTTTTCCACCTTTGGTTTAAATATCCTGGCAGAGTCTTCCGATTGG	1140
DB	1191	CATGTTCTACTGCTTTTCCACCTTTGGTTTAAACATATTTGGCAGAGCTCTCTGCTCGG	1250
QY	1141	TGATCGTGAATTTTACAAGGATTTGGTGAATGCCAAAACCTGTGGAAGATATTTGCGAGAT	1200
DB	1251	GGATCGTGAATTTTACAAGATTTGGTGAATGCAAAAAGTGTGGGAGATTAATCGGAAT	1310
QY	1201	GTGAATATGCGTGTTCACAAATGGATGATCGGCACCTATATTTTCCATGTTTAAGGCA	1260
DB	1311	GTGAATATGCGCTTTCATAAATGGATGTTGAGCTTCGACATATATACTTCCCGTCTGTCG	1370
QY	1261	CGGTCTACCAAGGCTGCTCTTTTAAATGCCCTTCTGGTTCTGCTTTATTTCCATGA	1320
DB	1371	CAAGATACCAAAACACTGCCCATTAATGCTTTCTTAGTCTCTGCGAGCTTTTCAATGA	1430
QY	1321	GCTGTGATTTGCTTTCTTGGCCACATATATCAAGTTGTGGCTTTTCGGTGGAAATATGTT	1380
DB	1431	ACTATGCATCGAGTCTCTGTGCTCTTTCAAGCTATGGCTTTTCTTGGGATTTGTT	1490
QY	1381	TCAGGTTCTTTGGCTGTGATCAATTAATCTGCAAAATAAATTCAGAAACTCAATGGT	1440
DB	1491	TCAGGTCGCTTTGGCTTCTACCAACTATCTACAGAAAGGTTTGG--CTCAACGGT	1547

Qy	1441	TGGAATATGATTTTTTGGTTTCATATTCAGTATACCTTGGTCAACCTATGTGTGTACTGCT	1500
Db	1548	GGGGAACATGATCTTCGGTTTCATCTTCTGCACTTTTCGGACAACCGATGTGTGTGCTTCT	1607
Qy	1501	ATATACCATGACTTGATGAATAGGAAGGCACAACTTGACTGGAAGCTACGG	1551
Db	1608	TTATTACCACCGCTGATGAACCGAAAAGGATCGATGTCATGAACAACACTG	1658
RESULT 10			
AF251794		1537 bp mRNA linear	PLN 16-APR-2000
LOCUS		Brassica napus putative diacylglycerol acyltransferase mRNA,	
DEFINITION		complete cds.	
ACCESSION		AF251794	
VERSION		AF251794.1	GI:7576940
KEYWORDS			
SOURCE		Brassica napus.	
ORGANISM		Brassica napus	
PLANT		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.	
REFERENCE		1 (bases 1 to 1537)	
AUTHORS		Brown,A.P., Schlierer,T.P. and Slabas,A.R.	
TITLE		Characterization of a putative diacylglycerol acyltransferase cDNA from Brassica napus embryo	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 1537)	
AUTHORS		Brown,A.P., Schlierer,T.P. and Slabas,A.R.	
TITLE		Direct Submission	
JOURNAL		Submitted (04-APR-2000) Biological Sciences, University of Durham,	
		South Road, Durham DH1 3LE, UK	
FEATURES		Location/Qualifiers	
source		1..1537	
		/organism="Brassica napus"	
		/cultivar="Commanche"	
		/db_xref="taxon:3708"	
		/dev_stage="embryo"	
		22..1527	
		/EC_number="2.3.1.20"	
		/codon_start=1	
		/product="putative diacylglycerol acyltransferase"	
		/protein_id="AA064065.1"	
		/db_xref="GI:7576941"	
		/translation="MEILDSGGVTPTENGADLTLRHKRSDSSNGLLPDSVTVS	
		LADVDVRVDSAGVEDTQKCANLAGENIRESGEAGGNVDVRYTTRSPVAHRRVRRES	
		LSDDATPKQSHAGLNLGVVLVAVNSRLIENLKYGLIRTDWFSTSLRDWPLF	
		MCGLSLIFPLAAFTVEKLVLOKCISEPVVILHLITTEVLVPVYVTLRCDASFLS	
		GVPLMLTCLIKWLVSYAHTNDIRTLANSKDKANPEVSYSLSLAYFMLPLATLC	
		YQSYSPRSCTKNGWVAQKLIITFGMFLIEQYINPIVRNSKHPKGLDILLYGVE	
		RVULKVYHVNYVLCMFYCFEHLMLIALLELCFGDREYKDMWNAKSGDYWRMMNI	
		PVHKMVRVHYFFCLRPPIKPAVILIALFLSAVHELGLIAPVCLFLKLAFTGLGIFM	
		VPLVFTINYLQERFGSMVGNIEFWFFCFIQGPQMCVLLYYHDLNRKRGKS"	
BASE COUNT		371 a 322 c 355 g 489 t	
ORIGIN			
Query Match		36.1%; Score 701.2; DB 8; Length 1537;	
Best Local Similarity		73.5%; Pred. No. 1.9e-151;	
Matches		933; Conservative 0; Mismatches 313; Indels 24; Gaps 2;	
Qy	264	TCAAATTCGCCATCCGTCCTTCCTCCGCTCCCGCTCACCGCAAGTCAAGCAAGTCCGCTCA	323
Db	272	TAAAGTACAGTATCGGCCGTCGGTTCACGCTCATCGAGGGTGGCGGAGAGTCCACTCA	331
Qy	324	GCTCCGACCTATTTTCCGTCAGAGTACCGCGGCCCTTTCAACCTTTGTATAGTAGTCC	383
Db	332	GCTCTGACGCCATCTTCAAACAGAGCCATGCTGGCACTATTCAACCTGCTGTAGTAGTTC	391
Qy	384	TTGTTGCTGTGAATAGCGGACTCATCATTTAGAGATTTAAATGAAGTATGGTTGATGATCA	443
Db	392	TTGTTGCTGTAAACAGTAGACTCATCATCGAAAAATCTCATGAAGTACGGTTGGTTGATCA	451
Qy	444	AATCTGGCTTTTGGTTTAGTTCAAAGTCATTGAGAGACTGCCCCCTTTTCATGTGTTGTC	503

Db	452	GAACGATTTCTGGTTAGTTCAACGCTCTCGGAGATTGGCCCTTTTCATGTGTGTC	511
QY	504	TTTCTCTTTGGTATTTCTTTCGCTGCTTTATAGTGGAGAGTTGGACACGGAAGT	563
Db	512	TCCTCCCTTTCAATCTTCTTGGCTGCTTTACCGTCGAGAAATAGTACTTCAGAAAT	571
QY	564	GATATACCCCAACCACTGTTGTTGACTTCATATATCAATACCTCAACTTCGCTTTTCT	623
Db	572	GCATATCTGAACCTGTGTCATCTTCTCATATTTATATACCATGACCGAGGCTTGT	631
QY	624	ATCCAGTTTATGTTATCTCAGGTGATCTGCTTTTGTATCAGGTGTCAGGTTAATGC	683
Db	632	ATCCAGTCTATGCTACTCTAAGGTGATTCGCCCTTCTTATCAGGTGTCACGTTGATGC	691
QY	684	TGTTTCTTGTGTATGTTTAAATTTGTTCTTATGTCATACATAAACTATGATATGA	743
Db	692	TCCTCACTTGTGATGTTGGCTGAAGTTGTTCTTACGCTCATATACTATGACATAA	751
QY	744	GAGCACTTACCAATTTAGTTGAAAGGGAGAACACTGCTCGATCTGAAACATGCACT	803
Db	752	GAACCTTACTAATTCATCTGATAAGG-----CCAAATCTCTGAAG	790
QY	804	ATCCTTACAAAGTTAAGCTTCAAGAGCTTGCATATTTCTGTTGCCCTACATTTATGTT	863
Db	791	TCCTCTACTATGTTAGCTTGAAGAGCTTGGGTATTTTCATGCTTGCCTCCACATTTGTT	850
QY	864	ACCAGCAAGCTATCCTCGCACACCTTATATTCGAAAGGTTGTTGTTTCGCCCACTTG	923
Db	851	ATCAGCCGAGCTATCCACGTTCTCCATGATATCCGGAAGGTTGGGTGGCTGTCATTTG	910
QY	924	TCAGCTGATATATTTACAGAGTATGGATTTATATAGACCAATATATTAATCCCA	983
Db	911	CAAGCTGATCATATTCACGTGATTCAGTGATTCAGGATTTATAATAGAGCAATATAAAATCCCTA	970
QY	984	TAGTACAAATTCACAGCATCTCTCAAGGAAACCTTCTTACGCCACCGAGAGGTTTC	1043
Db	971	TTGTTAGGAACCTCAAAACATCTTTGAAAGGGATCTCTTATACGGTGTGAAAGAGTGT	1030
QY	1044	TGAAGCTTTCTGTTCCAAATTTATATGTTGTTGCTCTGCAATGTTCTTATGCTTTTCCACC	1103
Db	1031	TGAAGCTTTTCAGTTCCAAATTTATACGTTGCTGCTGCAATGTTCTACTGCTTCTCCACC	1090
QY	1104	TTTGGTTAAATATCTCGGAGAGCTTCTTCGATTTGGTGTATCGTGAATTCACAGGATT	1163
Db	1091	TTTGGTTAAACATATTTGGCAGAGCTCTCTGCTTCGGGATCGTGAATTCACAAAGATT	1150
QY	1164	GTGGAATGCCAAACTGTCGAAGATTATTGGAGGATGGAATATGCTGTTTCACAAAAT	1223
Db	1151	GTGGAATGCCAAAGCTGGAGATTATTGGAGATGTTGGAATATGCTGTTTCAATAAT	1210
QY	1224	GGATGATCCGCCACCTATATTTTCCATGTTTAAAGGCAGCTCTACCAAAAGGCTGCTCT	1283
Db	1211	GGATGTTCCGACATGATATCTTTCCGTGCTTCCGAGAAATATACCGAAAGTACCGCTA	1270
QY	1284	TTTAAATGCCCTCTGCTTCTGCTTTTATTCATGAGCTGTGCAATGCTGTTTCTTCCG	1343
Db	1271	TTATCTTCTGCTTCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCT	1330
QY	1344	ACATATTTCAAGTTGTTGGGCTTTTGGTGAATATGTTTTCAGGTTCTCTTGGTCTGATCA	1403
Db	1331	GTCTCTTCAAACTATGGGCTTTCTTGGGATGATGTTTTCAGGTTCTTGGTATTTATCA	1390
QY	1404	CTAATTTATCGCAATAAATTCAGAACTCAATGTTTGGAAATATGATTTTTTGGTTCA	1463
Db	1391	CAAACTACCTACAAAGAAAGTTTGG--CTCCATGTTGGGAAACATGATATCTGTTT	1447
QY	1464	TATTCAGTATCTTGGTCAACCTATGTTGTTACTGCTTACTACCATGACTTGTGATGAATA	1523
Db	1448	CTCTTGCATTTTCGACAAACGATGTTGTTGCTTCTTTTATATCATCAGCACTTGTGAACC	1507
QY	1524	GGAAAGGCAA	1533
Db	1508	GCAAGGAAA	1517
RESULT 11			
AF164434			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Brassica napus.			
Brassica napus			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.			
1 (bases 1 to 1512)			
Nykiforuk C.L., Laroche, A. and Weselake, R.J.			
Isolation and Characterization of a cDNA Encoding a Second Putative			
Diacylglycerol Acyltransferase from a Microspore-derived Cell			
Suspension Culture of Brassica napus L. cv Jet Neuf (Accession No.			
AF164434). (PCR99-158)			
Plant Physiol. 121 (3), 1053 (1999)			
2 (bases 1 to 1512)			
Nykiforuk, C.L., Laroche, A. and Weselake, R.J.			
Direct Submission			
Submitted (30-JUN-1999) Chemistry and Biochemistry, University of			
Lethbridge, 4401 University Drive, Lethbridge, Alberta T1K 3M4,			
Canada			
Location/Qualifiers			
1..1512			
/organism="Brassica napus"			
/strain="cultivar Jet Neuf"			
/db_xref="taxon:3708"			
/note="microspore-derived cell suspension culture"			
1..1512			
/gene="Dgat1"			
1..1512			
/gene="Dgat1"			
/EC_number="2.3.1.20"			
/codon_start=1			
/product="putative diacylglycerol acyltransferase"			
/protein_id="AAD4536.1"			
/db_xref="GI:5579408"			
/translation="NAIDSGVAVPPTENGVAIDLRLHRRKSSDSSNGLSDTSPS			
DDVGAAARERDRVDSAAEEAGTANLAGDAETRESAGDVRFTYRPSVPAHRRTRE			
SPSSDAIFKOSHAGLNLVAVNSRLIENLMKVLIRTFDFWFSPTSLRDP			
LFMCCLSLVSPPLAAFTVEKVKVLOKEFISEPVAIIILHVIITHEVLVPPVTLRCDSP			
LSGVTMLLTCLVWLKLSYATSYDITLANSADKVDPEISYVLSKSLAYFMVAFT			
LCYQSPSPSPCIRKGVAROLAKLIVFTGLMGFTIEQYINPIVNSKHPKLDLLYA			
IERVLKSPNLYVWLCMFYCFHLLNLIAELLCTGDFEYKDWNNKSKVGDYWRW			
NMPVHKMVRHVPFLRIKIPKVPAILIAFLVSAVFHELCTAVPCRLFLNVAENGIM			
FOVPLVFTTNFLQERFGSMVGNMIFGSCIFGQPMCGLLYVHDLNMRKSGNS"			
BASE COUNT	344 a	343 c	363 g 462 t
ORIGIN			
Query Match	36.1%	Score 700.2;	DB 8; Length 1512;
Best Local Similarity	72.3%;	Pred. No. 3.3e-151;	
Matches	947;	Conservative	0; Mismatches 338; Indels 24; Gaps 2;
QY	235	CGAAACAAAGACACATGATTTCTCGTCTCTCAAAATTCGCTACCGTTCCTCCGTC	294
Db	228	CGAAATAGGAAATCCCGGAGGCGGATGTAAGGTTTAGGTATCGACCGTCCGCTCAGC	287
QY	295	TCACCGCAAGTGAAGAAAGTCCGCTCAGCTCCGACACTATTTCCGTCAGAGTCACGC	354
Db	288	TATCGGAGACGAGGAGAGTCTCTCAGCTCCGACGCTATCTTCAAAACAAAGCATGC	347
QY	355	GGGCTCTTCAACCTTTGTATAGTAGTCTCTGTTGTTGTAATAGCGGCTCATCATTTGA	414
Db	348	AGGATTTTCAACCTCTGTGTAGTCTTCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTT	407
QY	415	GAATTTAATGAAGTAGTGGTTGGTTGATCAAAATCTGGCTTTTGGTTTGAAGTCATT	474

Db	408	AAACCTCATAAGATATGGTATGATCAGAACTGATTTTGGTTTATGTTCTACATCCCTT	467
Qy	475	GAGAGACTGGCCCCCTTTTCATGTGTGTGCTTCTCTCTGTGTATTTCCCTTTCGCTGCCTT	534
Db	468	ACGAGACTGGCCGCTTTTCATGTGTGTCTTTCACCTTTTCGCTTCTCTTGGCTGCCTT	527
Qy	535	TATAGTGGAGAGTTGGCAACAACGGAAGTGATACCCGAACCAAGTGTGTGTGTACATCA	594
Db	528	CAGGGTCGAGAAAATGGTACTTTACAGAAAATCATATCTGAGCGCTGTGCGCATCATCTTCA	587
Qy	595	TATATCATTTACCTCAACTTCGCTTTTCTATCCAGTTTATGTTATTTCTCAGGTGTGATTC	654
Db	588	TGTATATTAACCATGACAGAGGTCTGTATCCAGTCTACGTCACACGTAGGTGTGATTC	647
Qy	655	TGCTTTTGTATCAGGTGTCCAGCTTAATGCTGTTTCTTGTGTGTATGTTTAAAAATTTGGT	714
Db	648	TGCTTCTTGTTCAGGTGTCCAGCTTGATGCTGCTACCTTGCATTTGTGTGGTGAAGTTTGGT	707
Qy	715	GTCTTATGCACATACAAACTATGATATGAGAGCACTTACCAAAATAGTTGAAAAGGAGA	774
Db	708	TTCTTACGCTCATACTAGCTACGACATAAGAACCCCTGGCCAAATTCAGCTGATAAGGTCGA	767
Qy	775	AGCACTGCTCGATCTCTGAACATGGACTATCCCTTACAACGTATAGCTTCAGAGCTTGGC	834
Db	768	TCC-----TGAATCTCCTACTATCTTACGTTGAAGAGCTTGGC	806
Qy	835	ATATTTCTCTGGTTGCCCTTACATATATGTTACGAGCAAGCTATCTCTCGCACACCTTATAT	894
Db	807	GTATTTTCATGGTGTCTCCACACTGTGTATTCAGCAAGCTATTCACGTTCTCCATGCTAT	866
Qy	895	TCGAAAGGGTGTGGTGTTCGCGCAACTGTCAAGCTGATATATTTACAGGAGTTATGGG	954
Db	867	CCGGAAGGGTGTGGTGTGCTCGTCAACTTGCAAAACGTGCTCATATTTCACTGGACTCATGG	926
Qy	955	ATTATTAATAGACCAATATATTAATCCCATAGTACAAAATTCACAGCATCTCTCAAGGG	1014
Db	927	ATTATTAATAGAGCAATATATAAATCCATATGTTTAGGAACTCAAAGCATCTCTGAAAGG	986
Qy	1015	AAACCTCTTTACGCCACCGAGAGAGTTCTGAAGCTTTCCTGCCAAATTTATATGTGTG	1074
Db	987	GGACCTTCTATATGCTATTGAAGAGTGTGNAAGCTTTCAGTTCCAAATCTATATGTGTG	1046
Qy	1075	GCCTCGATGTTCTATATGCTTTTCCACCTTTGGTTAAATATCTGCGAGAGCTTCTTCG	1134
Db	1047	GCCTCGATGTTCTACTGCTTCTTCCACCTTTGGTTAAACATATTTGCGAGAGCTCCCTCG	1106
Qy	1135	ATTTGGTGATCGTGAATTTCTACAAGGATTTGGTGGNAATGCCAAAACCTGTCGAAGATTAATG	1194
Db	1107	CTTTCGGGGACCGTGAATTTCTACAAAGATTTGGTGGAAATGCAAAAAGCGTTGGAGATTATG	1166
Qy	1195	GAGGATGTGGAATATGCGCTGTTACAAATGGATATCGCGCACCTATATTTTCCATGCTTT	1254
Db	1167	GAGAAATGTGGAATATGCCCTGTTTCACAAATGGATGTTTCGACATGTATACTTTCCGTCCT	1226
Qy	1255	AAGGCACGGTCTACCAAGGCTGCTGCTTTTAAATTCCTTCCTGGTTTCTGCTTTATTT	1314
Db	1227	GCGCATCAAGATACCAAAAGTACCCGCCATATCATTTGCTTTCTTAGCTCTCTCGAGCTTT	1286
Qy	1315	CCATGAGCTGTGCAATTCCTTTCGCGACATATTAAGTTGTGGGCTTTCGGTGGGAAT	1374
Db	1287	TCATGAGTTATGCATCGCAAGTTCCCTTGCGGCTCTCTTCAATCTATGGGCTTTCATGGGAAT	1346
Qy	1375	TATGTTTTCAGGTTCCCTTGGCTTCTGATCACTAATTAATCTGCAAAATAAATTCAGAACTC	1434
Db	1347	TATGTTTTCAGGTCCTTGGCTCTTATCACAACTTTTACAAAGAAAGGTTTGG---CTC	1403
Qy	1435	AATGGTTGGAAATATGATTTTTTGGTTCATATTCAGTATCTCTGGTCAAACTATGCTGT	1494
Db	1404	CATGGTGGGAACATGATCTTTGGTTCAGTCTTGTGCATTTTCGGACACCGATGTGTGG	1463
Qy	1495	ACTGCTTATACTACCATGACTTGAATAGGAAAGGCAAACTTGACTCA	1543

Db 1464 C G T T C T T A T T A C C A T G A C C T G A T G A C C G A A A G G A T C A C T G T C T T G A 1512

RESULT 12  
AF155224

LOCUS AF155224 Brassica napus putative diacylglycerol acyltransferase (DGAT2)  
DEFINITION mRNA, complete cds.  
ACCESSION AF155224  
VERSION AF155224.1 GI:5225381  
KEYWORDS  
SOURCE Brassica napus.  
ORGANISM Brassica napus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 1446)  
AUTHORS Nykiforuk,C.L., Laroche,A. and Weselake,R.J.  
TITLE Isolation and Sequence Analysis of a Novel cDNA Encoding a Putative Diacylglycerol Acyltransferase from a Microspore-derived Cell Suspension Culture of Brassica napus L. cv Jet Neuf (Accession No. AF155224). (PGR99-123)  
JOURNAL Plant Physiol. 120 (4), 1207 (1999)  
REFERENCE 2 (bases 1 to 1446)  
AUTHORS Nykiforuk,C.L., Laroche,A. and Weselake,R.J.  
TITLE Direct Submission  
JOURNAL Submitted (31-MAY-1999) Chemistry and Biochemistry, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta T1K 3M4,  
Canada

FEATURES  
source Location/Qualifiers  
1..1446  
/organism="Brassica napus"  
/cultivar="Jet Neuf"  
/db\_xref="taxon:3708"  
/note="microspore-derived cell suspension culture"  
gene 1..1446  
/gene="DGAT2"  
82..1107  
/gene="DGAT2"  
/EC\_number="2.3.1.20"  
/codon\_start=1  
/product="putative diacylglycerol acyltransferase"  
/protein\_id="AAD40881.1"  
/db\_xref="GI:5225382"  
translation="MCCLSLISFPLAAFTVEKLVLQKISEPWIFLHVITMTVEVLY  
PVTVLRCDSAFLSGDTLMILCTICIKLVWSVAHTNYDJRTLANSDKANPEVSYYYS  
LKSLAFLMLAPTLCYOPSFRSPRCIRKKWAKQFAKLVIPTGLMGFIISQIYNPIVRN  
SKHPLGDDLLYA IERVKLUSSPVIVWLCMFYCFPHLNLIAELLCFGDRDFYKDWM  
NAKSVDGYRWMMNPVHKVMRHYEPCLRIKVPYPAIIIAFLVASVFHELCIAPVC  
RLFNLFAMFGIMFOVPLVFTTNFLQRFGSMGVNMIFGSASCI FGQPMCGLLLYHDLM  
NRKGSMS"

BASE COUNT 398 a 291 c 278 g 479 t

ORIGIN

Query Match 31.2%; Score 605.6; DB 8; Length 1446;  
Best Local Similarity 73.1%; Pred. No. 2.5e-129;  
Matches 828; Conservative 0; Mismatches 279; Indels 25; Gaps 3;

Qy 413 GAGAAATTATGAAGTATGGTTGGTGATCAAACTCGCCTTTGGTTTAGTTCCAAAGTCA 472  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 2 GAAAATCTCATGAAGTACGGTTGGTTGTGATCAGAAGTCAATTTCTGGTTTAGTTCAAAGCTGC 61  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 473 TTGAGACATGGCCCCCTTTTCATGTGTGTGCTTCTCTCTTGTGGTATTTCCCTTTTCGCTGCC 532  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 62 CTGCGAGA -TTGCGCGCTTTTCATGTGTGTGCTCTCCCTTTCAACTCTTTCTTTGGCTGCC 120  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 533 TTTTAGTGGGAAGTTGGCCAACGAAGTGATATACCAGAACACAGTTGTGTGTACTT 592  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 121 TTTACCGTCGAGAATAATAGTACTTCAGAAATGCATATCTGAACCTGTGTGCATCTTTCTT 180  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 593 CATATAATCAATACCTCAACTTCGCTTTTCTATCCAGTTTTAGTTATTTCTCAGGTGTGAT 652  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 181 CATGTTATTATCACCATGACCGAGGCTTGTGATCCAGCTATGTCACTCTTAAGGTGTGAT 240  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||



Qy	653	TCTGCTTTTGATCAGGTGTCACGTTAATGCTGTTTCTTGTTGTTGATGGTAAAATTG	712
Db	241	TTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT	300
Qy	713	GTGCTTTATGCACATACAAAATATGATATGAGAGCAGTTACCAAATAGTTGAAAAGGGA	772
Db	301	TTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT TTTTTTTT	360
Qy	773	GAAGCAGTCTCGATACATCTGAACATGGACATATCCCTTACACAGTAAGCTTCAAGAGCTTG	832
Db	361	AACTCTGAAGTCT-----CCTACTATGTTAGCTTGAAGAGCTTG	399
Qy	833	GCATATTTCCCTGGTTGGCCCTACATTAATGTTACAGCCAAAGCTATCCTCGCACACCTTAT	892
Db	400	GCTTATTTTCATGCTTGCTGCCACATATGTTTATCAGCCAAAGCTATCCTCGCTTCCATGT	459
Qy	893	ATTGCAAGGGTGTGGTTGTTTTCGGCAACTGCTCAAGCTGATAAATATTTACAGGAGTATTG	952
Db	460	ATCCGGAAGGGTGGTGGCTGCTCAATTTGCAAAAGTGGTCATATTCACCTGGACTCAAG	519
Qy	953	GGATTTATAATAGACCAATATATTAATCCCATAGTACAAAATTCACAGCATCCTCTCAAG	1012
Db	520	GGATTTATAATAGACCAATATATAAATCCTATTTAGTAGAACTCAAGCATCCTCTCGAA	579
Qy	1013	GGAAACCTTCTTTAGCCACCGAGAGAGTCTGAAGCTTTCTGTTTCCAAATTTATATGTG	1072
Db	580	GGGAGACCTTCTATATGCTTATTTGAAAGAGTGTGAAGCTTTCAGTTTCCAAATCTATATGTG	639
Qy	1073	TGGCTCTGCATGTTCTATGCTTTTTCACACTTGGTTTAAATATCCTCGCAGAGCTTCTT	1132
Db	640	TGGCTCTGCATGTTCTACTGCTTCTTCACCTTTGGTTTAAACATATTTGGCAGAGCTCCTC	699
Qy	1133	CGATTTGGTGATCGTGAAATTTACAAGGATTTGGTGAATGCGCAAACTCTCGAAGATTAT	1192
Db	700	TGCTTCGGGGACCGTGAATTTCTACAAAGATTTGGTGAATGCAAAAAGCTTTGGAGATTAT	759
Qy	1193	TGGAGGATGGGAATATGCTGTTTCACAAATGGATGATCCGCGACCTATATTTTCCATGT	1252
Db	760	TGGAGAATGGGAATATGCTGTTTCACAAATGGATGTTTCGACATGATATCTTTCCGTGC	819
Qy	1253	TTAAGGCACGCTCTACCAAAGGCTGCTCTTTTAATTTGCTTCTCTGTTTCTGCTTTA	1312
Db	820	CTGGCATCAAGATACCAAAAGTACCCGCCATTAATTCATTTCTTAGTCTCTGCACTC	879
Qy	1313	TTCCATGAGCTGTGCATTTGCTTCTTTCGCACATATTTCAAGTTTGTGGCTTTTCGGTGA	1372
Db	880	TTTCATGAGTTATGATCGCGAGTCTTCTTCCGCTCTCTTCAATCATATGGCTTTTCATGGA	939
Qy	1373	ATTATGTTTCAGTTTCTTGGCTTGATCACTAATTAATCTGCAAAATAAATTCAGAAAC	1432
Db	940	ATTATGTTTCAGTCCCTTTGCTCTTATCAAACTTTTACAAGAAAGTTTGG---C	996
Qy	1433	TCAATGTTGGAAATATCATTTTTTGGTTCTATATTCAGTATCCTTGGTTCACCTATGTGT	1492
Db	997	TCCATGTTGGGAACATATGATCTTTGGTTTCAGTCTTCTTTCGCAATTTTCGGAACACCATGTGT	1056
Qy	1493	GTACTGCTATACTACCATGACTTGCATGAATAGGAAAGGCAAACTTTGACTGAA	1544
Db	1057	GGGCTCTTTTATACCATGACCTGATGACCGCAAGGATCCATGCTCTGAA	1108
RESULT 13	AX090350	Sequence 43 from Patent WO0116308.	
LOCUS	AX090350	234 bp DNA linear	PAT 21-MAR-2001
DEFINITION	Sequence 43 from Patent WO0116308.		
ACCESSION	AX090350		
VERSION	AX090350.1	GI:13444211	
KEYWORDS	soybean.		
SOURCE	Glycine max		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		

Db 61 GATGATCGCCACCTATATTTTCCATGTTTAAAGGACGGTATACAAAGGCGTGTCTCT 120  
Qy 1285 TTTAATTCCTTCTCGTTCTGTTTATTCATGAGCTGTGCATTTGCTTCTTGTGCA 1344  
Db 121 TTTAATTCCTTCTCGTTCTGTTTATTCATGAGCTGTGCATTTGCTTCTTGTGCA 179  
Qy 1345 CATATTCAGTTGTTGGGTTTTCGGTGGAAATATGTTTCAGGTTTCTTGGTCTTGAATC 1404  
Db 180 ACATATTCAGTTGTTGGGTTTTCGGTGGAAATATGTTTCAGGTTTCTTGGTCTTGAATC 239  
Qy 1405 TAATTAATTCAGTTGTTGGGTTTTCGGTGGAAATATGTTTCAGGTTTCTTGGTCTTGAATC 1439  
Db 240 AATTNTNGGCAAAATAATTCNNGAACCCCGGG 274

RESULT 15  
AX090351  
LOCUS AX090351 267 bp DNA linear PAT 21-MAR-2001  
DEFINITION Sequence 44 from Patent WO0116308.  
ACCESSION AX090351  
VERSION AX090351.1 GI:13444212  
KEYWORDS  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 267)  
Lassner, M. and van Eenennaam, A.  
TITLE Plant sterol acyltransferases  
JOURNAL Patent: WO 0116308-A 44 08-MAR-2001;  
MONSANTO COMPANY (US)  
FEATURES  
source Location/Qualifiers  
1..267  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
BASE COUNT 70 a 52 c 57 g 88 t  
ORIGIN

Query Match 9.68; Score 187.2; DB 6; Length 267;  
Best Local Similarity 83.68; Pred. No. 1.2e-32;  
Matches 224; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

Qy 654 CTGCTTTTGTATCAGGTGTACGTTAATGCTGTTTCTTGTGTTGTTGTTAAATTTGG 713  
Db 1 CTGCTTTTGTATCTGGTGTACGTTGAGCTATTAACTTGCATTGTGTTGTTAAATTTGG 60  
Qy 714 TGTCTTATGCATACAACTATGATATGAGGACCTTACCATAATTTAGTTGAAAGGGAG 773  
Db 61 TGTCTATATGCATACAACTATGATATGAGGACCTTACTGTTTCGAATGAAAGGGAG 120  
Qy 774 AAGCACTCTCGTACTCTGAACATGAGTATCTTACAACTGAGCTTCAAGAGCTTGG 833  
Db 121 AAACATACCCCAACTTTTG-ATATGGAGTATCCGTACACTGTGACCTTCAGGATTTGG 179  
Qy 834 CATATTTCTGTTGCCCTACATTTATGTTACCGCAAGCTATCTCGCACACCTTATA 893  
Db 180 CATACTTCATGTTGCTCCTACATTTATGCTATCAGACAAGCTATCTCGCACACCTTCAG 239  
Qy 894 TTCGAAGGGTTGGTTGTTTCGCCAACT 921  
Db 240 TTCGAAGGGTTGGTTGTTTCGCCAACT 267

Search completed: February 21, 2003, 17:47:13  
Job time : 5051 secs